Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1778 (2008) 1897–1929
Biochimica et Biophysica Acta BBA
www.elsevier.com/locate/bbamem

Review

The prokaryotic complex iron–sulfur molybdoenzyme family

Richard A. Rothery, Gregory J. Workun, Joel H. Weiner *

Membrane Protein Research Group, Department of Biochemistry, 474 Medical Sciences Building, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received 20 May 2007; received in revised form 17 August 2007; accepted 2 September 2007
Available online 18 September 2007

Abstract

Bacterial genomes encode an extensive range of respiratory enzymes that enable respiratory metabolism with a diverse group of reducing and oxidizing substrates under both aerobic and anaerobic growth conditions. An important class of enzymes that contributes to this broad diversity is the complex iron–sulfur molybdoenzyme (CISM) family. The architecture of this class comprises the following subunits. (i) A molybdo-*bis* (pyranopterin guanine dinucleotide) (Mo-*bis*PGD) cofactor-containing catalytic subunit that also contains a cubane [Fe-S] cluster (FS0). (ii) A four-cluster protein (FCP) subunit that contains 4 cubane [Fe-S] clusters (FS1–FS4). (iii) A membrane anchor protein (MAP) subunit which anchors the catalytic and FCP subunits to the cytoplasmic membrane. In this review, we define the CISM family of enzymes on the basis of emerging structural and bioinformatic data, and show that the catalytic and FCP subunit architectures appear in a wide range of bacterial redox enzymes. We evaluate evolutionary events involving genes encoding the CISM catalytic subunit that resulted in the emergence of the complex I (NADH:ubiquinone oxidoreductase) Nqo3/NuoG subunit architecture. We also trace a series of evolutionary events leading from a primordial Cys-containing peptide to the FCP architecture. Finally, many of the CISM archetypes and related enzymes rely on the *tat* translocon to transport fully folded monomeric or dimeric subunits across the cytoplasmic membrane. We have used genome sequence data to establish that there is a bias against the presence of soluble periplasmic molybdoenzymes in bacteria lacking an outer membrane.

Keywords: Nitrate reductase; Formate dehydrogenase; Iron sulfur; Molybdenum cofactor; Structural biology; Cell envelope

Contents

Abbreviations: CISM, complex iron–sulfur molybdoenzyme; DmsABC, *E. coli* DMSO reductase; DorA, soluble *Rhodobacter* DMSO reductase; E_m, midpoint potential; EPR, electron paramagnetic resonance; ETR, electron transfer relay; EXAFS, extended X-ray absorption fine structure; FCP, four cluster protein; FHL, formate hydrogen lyase; FdnGHI, *E. coli* formate dehydrogenase N; MAP, membrane anchor protein; Mo-*bis*PGD, molybdo-*bis*(pyranopterin guanine dinucleotide); NarGHI, *E. coli* nitrate reductase A; PDB, protein data bank; rmsd, root–mean–square deviation; Sec, selenocysteine; SSM, secondary structure matching; TCP, three-cluster protein; TM, transmembrane; TorA, periplasmic TMAO reductase

* Corresponding author. Tel.: +1 780 492 2761; fax: +1 780 492 0886.
E-mail address: Joel.Weiner@UAlberta.ca (J.H. Weiner).

0005-2736/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2007.09.002

1. Archetypal enzymes of modular design

Respiratory electron transfer chain enzymes are absolutely essential for the exploitation of the diversity of ecological niches on Earth. Bacteria are found in all of these niches and are able to exploit a vast array of growth conditions, including highly oxidizing and reducing conditions, extremes of high and low temperatures, alkalinity and acidity, and even high levels of ionizing radiation. An essential factor in the success of the plethora of bacterial species in comparison with eukaryotes is their broad metabolic diversity. This is based on genomes that encode enzymes able to catalyze a remarkable range of substrate interconversions. An important subset of these involve redox reactions, and many of the enzymes responsible are constituents of highly flexible respiratory chains that have core components embedded in the cytoplasmic membrane. Importantly, many of these enzymes allow growth on a variety of reducing and oxidizing substrates under both aerobic and anaerobic growth conditions [1–4]. This allows exploitation of environments containing a range of potential terminal electron donors and acceptors. A group of enzymes that play a pivotal role in supporting this respiratory diversity is the complex iron–sulfur molybdoenzyme (CISM) family. This review focuses on describing this family and its functional and evolutionary relationships with a plethora of redox enzyme systems.

An archetypal CISM has an overall structure comprising three subunits with the following general characteristics (Fig. 1) [5–9]. (i) A molybdo-*bis*(pyranopterin guanine dinucleotide) (Mo-*bis*PGD, Table 1) containing catalytic subunit that also contains a [4Fe-4S] cluster known as FS0. (ii) An electron transfer subunit that typically contains four cubane [Fe-S] clusters known as FS1 to FS4 in sequence leading away from the Mo-*bis*PGD cofactor. This subunit is referred to as the four-cluster protein (FCP) subunit. Because they are able to catalyze soluble substrate interconversions in the presence artificial electron donors or acceptors, the catalytic and FCP subunits are often referred to as the catalytic dimer. (iii) A membrane anchor protein (MAP) that may or may not contain heme. The MAP subunit functions both to anchor the catalytic and FCP subunits to the cytoplasmic membrane and to provide the CISM with a redox-active quinol/quinone-binding site (Q-site; Fig. 1A). The [4Fe-4S] cluster of the catalytic subunit (FS0) and the four

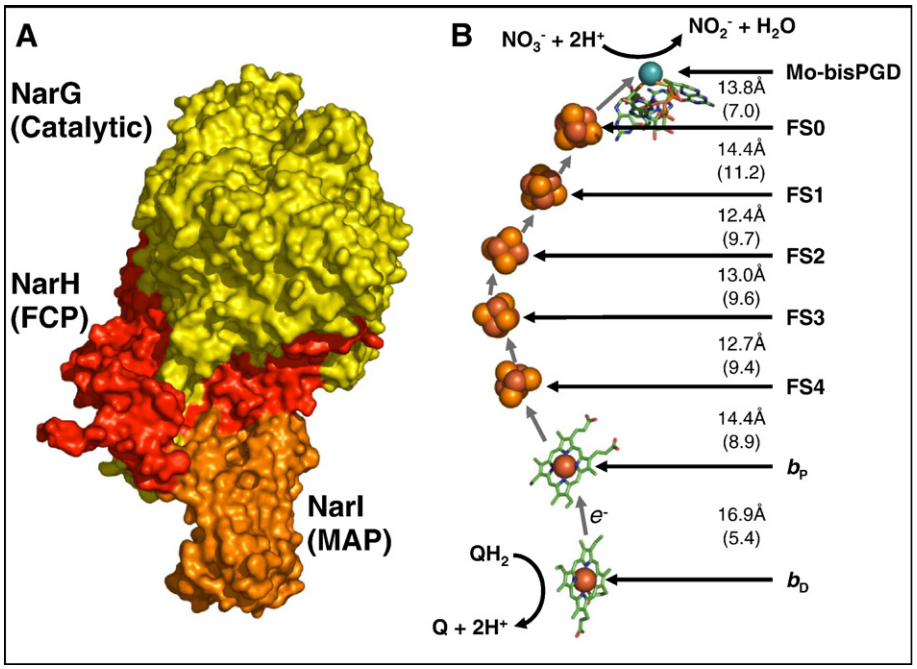

Fig. 1. Nitrate reductase A: an archetypal CISM. The three subunits of NarGHI (A) provide a molecular scaffold for two substrate binding sites and for the electron transferring relay that connects them (B). Structural representations of the PDB file 1Q16 [9] were rendered using the PYMOL molecular visualization package (DeLano Scientific LLC, Palo Alto, CA).

cubane clusters of the FCP (FS1–FS4) comprise an electron transfer relay (ETR) that shuttles electrons between the site of oxo-transfer and/or redox catalysis at the Mo-*bis*PGD active site and the site of quinone/quinol interconversion at the membrane-intrinsic Q-site (Fig. 1B). Thus, the archetypal CISM couples the unique redox and catalytic properties of the Mo-*bis*PGD cofactor to the versatile electron sink provided by the membrane-intrinsic quinone pool (the Q-pool). Importantly, in the context of this review, the overall structures of members of the CISM family are of a modular design, elements of which appear in a wide range of enzymes that are responsible for multiple levels of diversity within bacterial metabolomes.

One intriguing aspect of the CISM family is the variability of the orientation of the catalytic and FCP subunits with respect to the cytoplasmic membrane. This is dependent on the presence of a twin arginine translocase leader (*tat* leader) sequence at the N-terminus of the catalytic subunit that can direct the catalytic and FCP subunits to the periplasmic compartment via the *tat* translocon [10–12]. In this context, the FCP subunit is referred to as a "passenger protein". In CISM enzymes that lack a *tat* leader, the catalytic and FCP subunits are targeted to the inner surface of the cytoplasmic membrane. One consequence of this is that a complete respiratory chain can comprise two members of the CISM family with their catalytic dimers having opposite orientations with respect to the cytoplasmic membrane that are redox-coupled by the Q-pool. An example of this is the proton translocating respiratory chain comprising the archetypal *Escherichia coli* CISMs formate dehydrogenase N (FdnGHI) and nitrate reductase A (NarGHI) (Fig. 2). These enzymes have had their structures solved by X-ray crystallography [5,8,9], and in combination they embody the concept of proton translocation by a redox loop mechanism as originally proposed by Peter Mitchell [13]. Proton translocation is achieved via the scalar distribution of proton releasing and proton-consuming redox reactions across the energy-conserving membrane. For

Table 1
Mononuclear molybdenum cofactor nomenclature

Cofactor abbreviation	Generic identification	Comments
MoCo	Molybdenum cofactor	Generic term for the mononuclear molybdenum cofactor.
Mo-MPT	Mo-molybdopterin	The original structure proposed by Johnson and coworkers [29,30]. Not yet observed in an enzyme structure.
Mo-PPT	Mo-pyranopterin	The most common eukaryotic form containing a tricyclic pterin, for example, in sulfite oxidase.
Mo-*bis*PPT	Mo-*bis*Pyranopterin	A form found in archeal aldehyde oxidoreductase, often with tungsten substituted for the Mo [184].
Mo-*bis*PGD	Mo-*bis*(pyranopterin guanine dinucleotide)	The common prokaryotic form originally revealed by the structure of *Rhodobacter* DMSO reductase [14,15] and many other enzymes. Formerly referred to as Mo-*bis*MGD (Mo-*bis* (molybdopterin guanine dinucleotide).
PGD-Mo-MGD	(Molybdopterin guanine dinucleotide)-Mo-PGD	The form found in NarGHI [9]. The MGD pterin may exist in an equilibrium with the PGD form. For the sake of brevity, it will be referred to as Mo-*bis*PGD herein.

The structures of a large number of enzymes containing various forms of the mononuclear molybdenum cofactor molybdoenzyme structures have become available in the last decade [35]. These structures provide an opportunity to clarify the nomenclature used to describe the various forms of the mononuclear molybdenum cofactor.

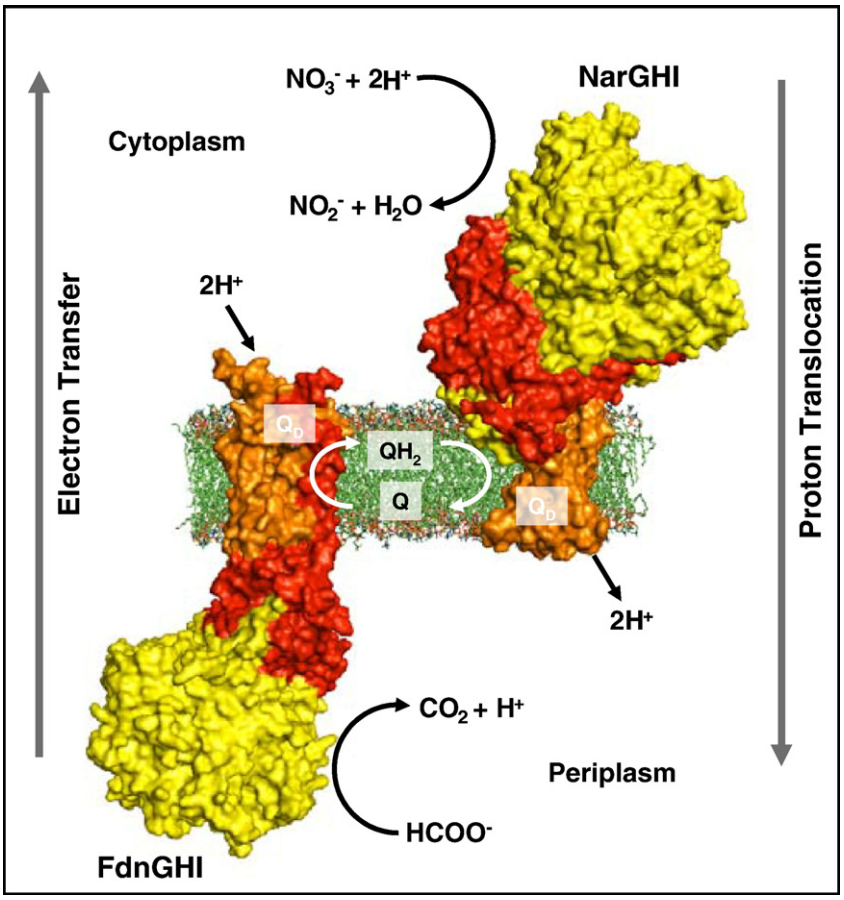

Fig. 2. Proton translocation by a redox loop mechanism. The formate:nitrate respiratory chain translocates protons via a mechanism involving scalar redox chemistry. The figure was prepared as described for Fig. 1 using the PDB files 1Q16 (NarGHI) and 1KQF (FdnGHI [8]).

example, in the case of the quinone redox chemistry catalyzed by these enzymes, FdnGHI catalyzes a proton-consuming quinone reduction on the cytoplasmic side of the membrane, whereas NarGHI catalyzes a proton-releasing quinol oxidation on the periplasmic side of the membrane. In the case of the soluble substrate redox chemistry, formate is oxidized in a proton-yielding reaction at a periplasmically localized active site, whereas nitrate is reduced in a proton-consuming reaction at a cytoplasmically localized active site (Fig. 2). These reactions clearly result in net proton translocation from the cytoplasm to the periplasm, contributing to the proton electrochemical potential across the cytoplasmic membrane.

2. Modularity of the CISM archetypes

Archetypes of the CISM family include enzymes that reduce substrates such as nitrate, *S*-oxides, *N*-oxides, thiosulfate, tetrathionate, and polysulfide, and enzymes that oxidize formate in a wide range of species (Table 2A), contributing significantly to the remarkable metabolic diversity of the bacterial and archaeal domains of life (Section 6). Each of these enzymes catalyzes a redox reaction at a Mo-*bis*PGD cofactor (Fig. 3, Table 1) coordinated by the catalytic subunit. Emerging structural data have revealed intriguing differences between the catalytic, FCP, and MAP subunits of these enzymes.

2.1. Structural diversity of the catalytic subunits

The catalytic subunit provides three functionalities to the archetypal CISM: (i) a molecular scaffold for the complex and highly labile Mo-*bis*PGD cofactor (Fig. 3); (ii) a substrate binding funnel allowing facile substrate entry and egress to and from the site of catalysis at the crucial Mo atom; and (iii) the segment of the ETR incorporating the FS0 [4Fe-4S] cluster leading to or from the [Fe-S] clusters of the FCP.

The architecture of the molecular scaffold of the CISM catalytic subunit is similar to that of the periplasmic DMSO reductases (DorA) of *Rhodobacter capsulatus* and *R. sphaeroides* [14,15]. The structure of the latter can be considered as the core structure of the catalytic subunits of the entire CISM family. DorA is arranged in four domains (I–IV) surrounding the Mo-*bis*PGD binding pocket [16]. Domains I, II, and III surround the active site funnel that leads to the site of DMSO reduction at the Mo atom. Domain IV lies beneath the Mo-*bis*PGD cofactor. Typically, the catalytic subunits of the CISM family have the following additional structural and functional features: (i) an additional domain of variable size, Domain V, that plays an essential role in defining the substrate binding funnel; (ii) an N-terminal Cys group located in Domain I that provides coordination to the [Fe-S] cluster (FS0) of the ETR that is closest to the Mo-*bis*PGD cofactor. This cluster and

Table 2
Members and close relatives of the CISM family

Abbreviation of system	Name of system	Example organism	Subunit characteristics			
			Catalytic	FCP	MAP	Other
A. Archetypes						
NarGHI	Nitrate reductase A	*E. coli*	1246 NarG	512 NarH	225 NarI	MAP has 5 TMS
FdnGHI	Formate dehydrogenase	*E. coli*	982 FdnG●	294 FdnH	217 FdnI	MAP related to CybH (4 TMS)
PhsABC	Thiosulfate reductase	*S. typhimurium*	728 PhsA●	192 PhsB	254 PhsC	MAP related to CybH (5 TMS)
DmsABC	DMSO reductase	*E. coli*	814 DmsA●	204 DmsB	287 DmsC	MAP related to NrfD and PsrC (8 TMS)
PsrABC	Polysulfide reductase	*W. succinogenes*	763 PsrA●	191 PsrB	317 PsrC	MAP related to NrfD and DmsC (8 TMS)
TtrABC	Tetrathionate reductase	*B. parapertussis*	1007 TtrA●	257 TtrB	344 TtrC	MAP weakly related NrfD (9 TMS)
B. Related—contains one or more archetype subunits						
(i) CISM dimer—contains catalytic and FCP subunits only						
DdhABC	Dimethyl sulfide dehydrogenase	*R. sulfidophilum*	882 DdhA●	325 DdhB	≡	240 DdhC ◆—monoheme cytochrome *b*
SerABC	Selenate reductase	*T. selenatis*	883 SerA●	326 SerB	≡	211 SerC ◆—monoheme cytochrome *b*
ClrABC	Chlorate reductase	*I. dechloratans*	882 ClrA●	328 ClrB	≡	212 ClrC ◆
EbdABC	Ethylbenzene dehydrogenase	*Azoarcus* sp. EB1	931 EbdA●	352 EbdB		214 EbdC—monoheme cytochrome *b* (no Sec leader)
PgtLS	Pyrogallol–phloroglucinol transhydroxylase	*P. acidigallici*	874 PgtL	274 PgtS		
FdhAB	Formate dehydrogenase	*D. gigas*	977 FdhA●	214 FdhB	≡	Interacts with a cytochrome c_3
(ii) Catalytic—contains only the catalytic subunit						
TorAC	TMAO reductase	*E. coli*	809 TorA●	↔	↔	390 TorC—pentaheme cytochrome *c*
NapAB	Periplasmic nitrate reductase	*E. coli*	797 NapA●	≡	≡	≡ 122 NapB ◆—diheme cytochrome *c*
BisC	Biotin-*d*-sulfoxide reductase	*E. coli*	777 BisC	↔	↔	↔ 200 NapC—tetraheme cytochrome *c*
AoxAB	Arsenite oxidase	*A. faecalis*	825 AoxB	≡	≡	133 AoxA●—Rieske-type [2Fe-2S] cluster interacts with azurin
Nqo1-15	Complex I (NADH-UQ oxidoreductase)	*T. thermophilus*	783 Nqo3	≡	≡	In multisubunit complex with ~14 other subunits, lacks Mo-*bis*PGD and FS0
(iii) FCP and anchor—contain FCP and membrane anchor subunits only						
HyaABC (MbhSL-CybH)	Hydrogenase-1	*E. coli*		327 HyaA● (MbhS)[a]	234 HyaC (CybH)	597 HyaB—hydrogenase 1 Ni-containing subunit (MbhL)
NrfABCD	Nitrite reductase	*E. coli*		196 NrfC●	318 NrfD related to DmsC/PsrC	156 NrfB ◆—pentaheme cytochrome *c*; ↔ 452 NrfA ◆—nitrite reducing cytochrome *c* (pentaheme)
HmcBC-Hmwc (Hmc23-Hmwc) PhfLS (HydAB)	Hydrogenase system (electron donor system)	*D. vulgaris*		343 HmcB● (Hmc2)	388 HmcC (Hmc3)	↔ 328 Hmwc ◆ ↔ 421 PhfL (HydA) ≡ 89 PhfS● (HydB)

●—Precursor has an *N*-terminal *tat* leader sequence that is cleaved in the mature form of the protein.
◆—Precursor has an *N*-terminal *sec* leader sequence that is cleaved in the mature form of the protein.
↔—Denotes transient interaction.
≡—Denotes stable subunit–subunit interaction within multisubunit complexes.
[a] Not a true FCP, contains 2 [4Fe-4S] clusters and one [3Fe-4S] cluster.

its associated Cys residues are absent in the periplasmic DMSO and TMAO reductases (DorA and TorA).

An important feature of the domains of the CISM catalytic subunit is their largely non-contiguous location within the primary protein sequence (Fig. 4A). This complex architecture is evolutionarily derived from a currently untraceable series of gene insertion/duplication events. Secondary structure matching (SSM) analyses [17] can provide insights into the differences between DorA-like subunits and archetypal CISM catalytic subunits such as NarG or FdnG. Fig. 4B shows the results of an SSM analysis comparing NarG with DorA and *E. coli* FdnG. Structural overlap between NarG and DorA encompasses 602 residues with an rmsd of 2.75 Å, representing 45% and 77% of the total residues, respectively. Comparison of the structures of NarG and FdnG reveals that 621 residues align with an rmsd of 2.37 Å, representing 50% and 63% of the total residues, respectively. Much of the sequence of NarG that does not align structurally with that of DorA is located within the non-contiguous Domain V and plays a significant role in defining the much narrower substrate binding funnel of NarG in comparison with that of DorA (Fig. 5). As is also the case for the FCP subunit (NarH, see Section 2.2), much of this sequence, both that of Domain V and elsewhere, is located on the surface of the heterotrimeric holoenzyme [9].

Given the large amount of protein synthesis involved in generating the additional surface-localized subdomains of NarG and NarH, it is unlikely that these arose merely by chance. In the case of the extra sequence located in Domain V of NarG, it is possible that its role is to modulate substrate specificity towards nitrate and the very small number of oxyanions that have been identified as substrates [18–20], including nitrate, bromate,

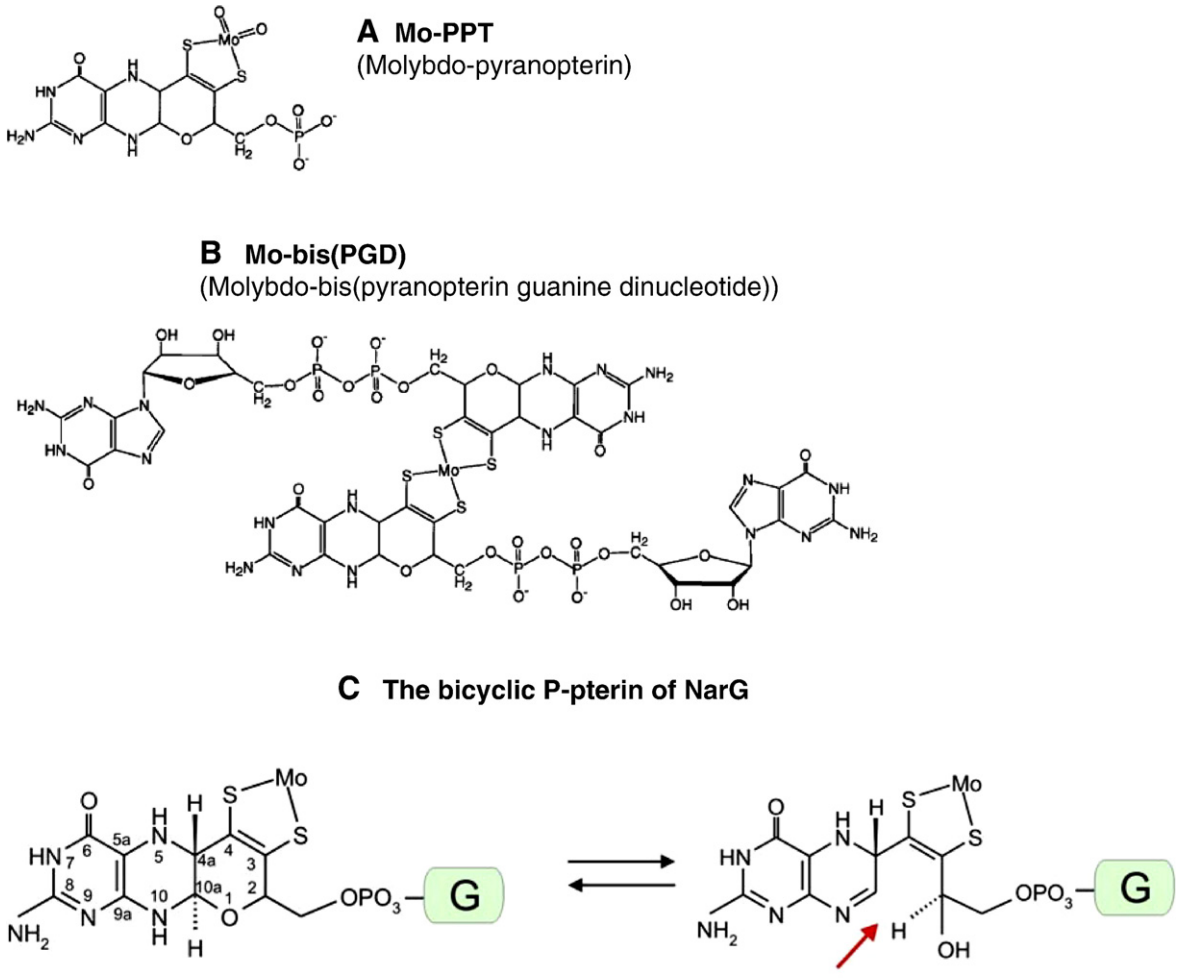

Fig. 3. Common forms of the mononuclear molybdenum cofactor. Two forms of the molybdenum cofactor found in *Escherichia coli* (A, B) and the bicyclic form of the Q-pterin found in the catalytic subunit of NarG (C). The "G" in the green box represents the guanine nucleotide.

chlorate, selenate and tellurite. It is notable that DorA lacks Domain V, has a fairly open substrate binding funnel (Fig. 5), and displays a relatively broad substrate specificity towards *S*- and *N*-oxides [21–23]. Likewise, although its structure is not yet known, the membrane-bound CISM DMSO reductase from *E. coli* (DmsABC) is also predicted to lack Domain V and it retains the broad substrate specificity of DorA [24,25]. The additional surface-localized subdomains of NarG (and NarH) could be an evolutionary adaptation to the reactive nitrogen species that are intermediates of the global nitrogen cycle [26,27]. It is well known, for example, that nitric oxide can seriously deplete the [Fe-S] content of mitochondrial respiratory chain complexes [28].

NarG and FdnG each contain two antiparallel pterins, each covalently linked by a phosphodiester bond to a guanine nucleotide (Table 1; Fig. 3) [7–9]. The pterin proximal to the FS0 [4Fe-4S] cluster is termed the P-pterin, and the one distal to it is termed the Q-pterin. In the case of FdnG [8], both pterins have a tricyclic pyranopterin structure (Fig. 3) that is a modification of the molybdopterin structure originally proposed by Rajagopalan and coworkers [29–31]. This is the form of the cofactor identified in the structures of DorA [14,15], the periplasmic nitrate reductases (NapA) from *R. sphaeroides* [32] and *Desulfovibrio desulfuricans* [33], *E. coli* formate dehydrogenase H (FdhF) [34], and in a range of molybdoenzymes containing the *bis*-guanine dinucleotide form of the cofactor. Thus, in the vast majority of enzymes structurally characterized to date, the cofactor is in the form of a molybdo-*bis*(pyranopterin guanine dinucleotide) (Mo-*bis*PGD) [35–37] (Fig. 3). This emerging homogeneity of chemical structure has been confounded by structural analysis of the cofactor in NarGHI, which reveals it to have a PGD-Mo-MGD structure (Table 1) [9]. In this case, the P-pterin is a pyranopterin, and the Q-pterin is a molybdopterin. The molybdopterin structure of the Q-pterin appears to be stabilized by hydrogen bonds between the free hydroxyl of the open pyran ring and the absolutely conserved NarG-S719 and NarG-H1163 residues. It has been speculated that pyran ring opening and closing may play a role in catalysis, perhaps aiding proton translocation to the site of nitrate reduction at the redox-active Mo atom [9,37,38]. It has also been noted that the cofactor in ethylbenzene dehydrogenase (EbdABC) from *Aromatoleum aromaticum* also has a PGD-Mo-MGD structure [39], and in this case the residues stabilizing the open pyran ring are an Arg (EbdA-R612) and a His (EbdA-

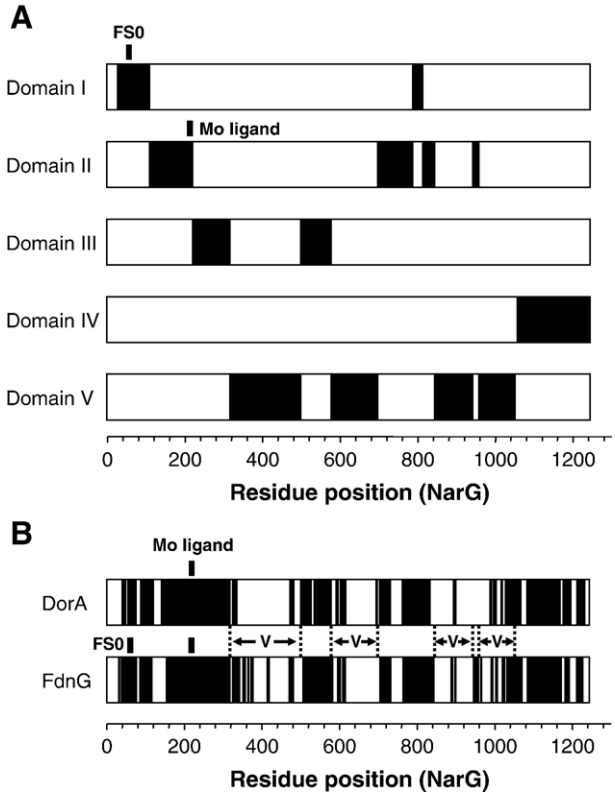

Fig. 4. (A) Location of the five domains of NarG in its primary sequence. The positions of the residues coordinating FS0 and the Mo of the Mo-*bis*PGD cofactor are as indicated in black. (B) Structural overlap of NarG with DorA and FdnG. Segments shown in black represent the sequence segments of NarG that overlap structurally with DorA (PDB code 1DMS [14]) or FdnG (PDB code 1KQF [8]). Segments shown in white are absent from DorA or FdnG. Note the almost complete absence of Domain V from DorA and its partial absence from FdnG. Secondary structure matching was performed using the EBI SSM server at (http://www.ebi.ac.uk) [17].

H855). EbdABC is not an archetypal CISM and is discussed in more detail in Section 3.1.

The Mo coordination environment of the CISM catalytic subunits comprises the four dithiolene sulfurs of the two pterins, a protein–Mo ligand that is typically a Ser, selenocysteine (Sec), or Cys, and a hydroxyl or an oxo group [35,36,38]. The protein–Mo ligand of NarG is provided by the carboxylate sidechain of an Asp residue (NarG-D222). This appears to be able to coordinate the Mo via both carboxylate oxygens in a bidentate manner [9], or via a single carboxylate oxygen and a separate oxo group [7]. Because of the limitations of X-ray crystallographic analyses of large multisubunit enzymes such as NarGHI, clarification of the Mo coordination environment will require complementary analyses using techniques such as EXAFS (extended X-ray absorption fine structure) and EPR (electron paramagnetic resonance) spectroscopies of the wild-type and mutant enzymes. The coordination environments of other Mo-*bis*PGD containing enzymes have been extensively reviewed and will not be covered herein [35,36,38].

Domain I sequences of the CISM catalytic subunits generally contain either four Cys residues or three Cys residues and one His residue [40–42]. Original speculation that in each of the CISM enzymes these residues coordinate a [4Fe-4S] cluster (FS0) was vindicated by the emergence of the structures of NarGHI and FdnGHI. In the case of NarGHI, EPR studies indicated that the reduced FS0 cluster exists in an $S=3/2$ (high spin) ground state and has a midpoint potential (E_m) of approximately −55 mV [43]. As a result of its unusual spin state, its EPR spectrum is observed between $g=5.02$ and 5.56, rather than comprising features clustered around approximately $g=1.94$, as is the case with typical [4Fe-4S] clusters with an $S=1/2$ (low spin) ground state.

Of the known CISM enzymes, only two, NarGHI and DmsABC, have been subjected to intense scrutiny by EPR spectroscopy [41–63]. In the case of DmsABC, no direct EPR evidence has been obtained for the presence of an FS0 cluster, and, as yet, there is no structure available. However, there is a strong EPR-detectable spin–spin interaction between the Mo atom of the cofactor and a nearby paramagnetic center which has an E_m of approximately −140 mV [57,62]. Originally, it was suggested that this interaction occurs between the Mo and one of the four [4Fe-4S] clusters of DmsB, whose E_m values are approximately −240 mV (FS1), −330 mV (FS2), −120 mV (FS3), and −50 mV (FS4) (Fig. 6), with FS3 ($E_m=-120$ mV) being the prime candidate for being the interacting cluster [57,61]. However, the emergence of the structures of FdnGHI and NarGHI renders a Mo-FS3 interaction highly implausible because of the large intercenter distance involved (~40 Å in the case of the structure of FdnGHI). Re-examination of the data in

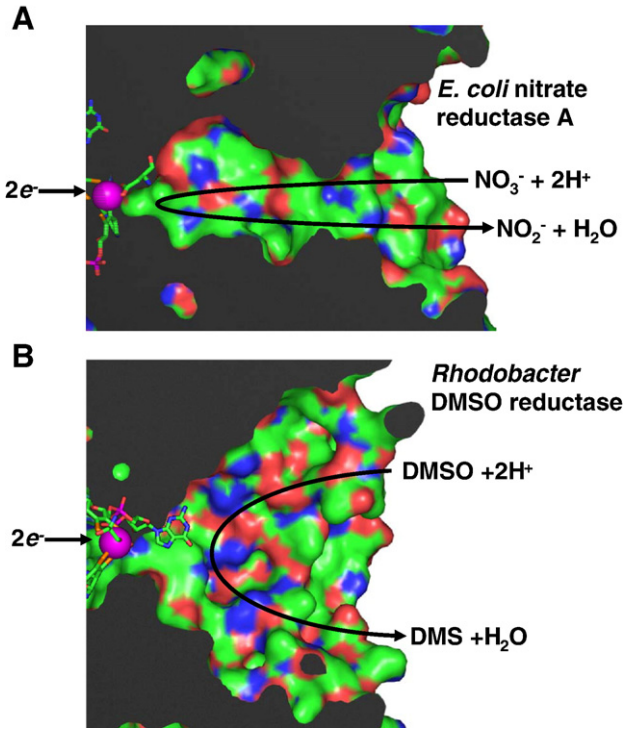

Fig. 5. Comparison of the substrate binding funnels of *Escherichia coli* NarGHI (A) and Rhodobacter DorA (B). Note that in both cases the Mo-*bis*PGD cofactor is shown for reference purposes only and that the guanine nucleotide visible in panel B is above the observed substrate funnel. Images were rendered from PDB files 1Q16 [9] and 1DMS [14] using the PYMOL molecular visualization package.

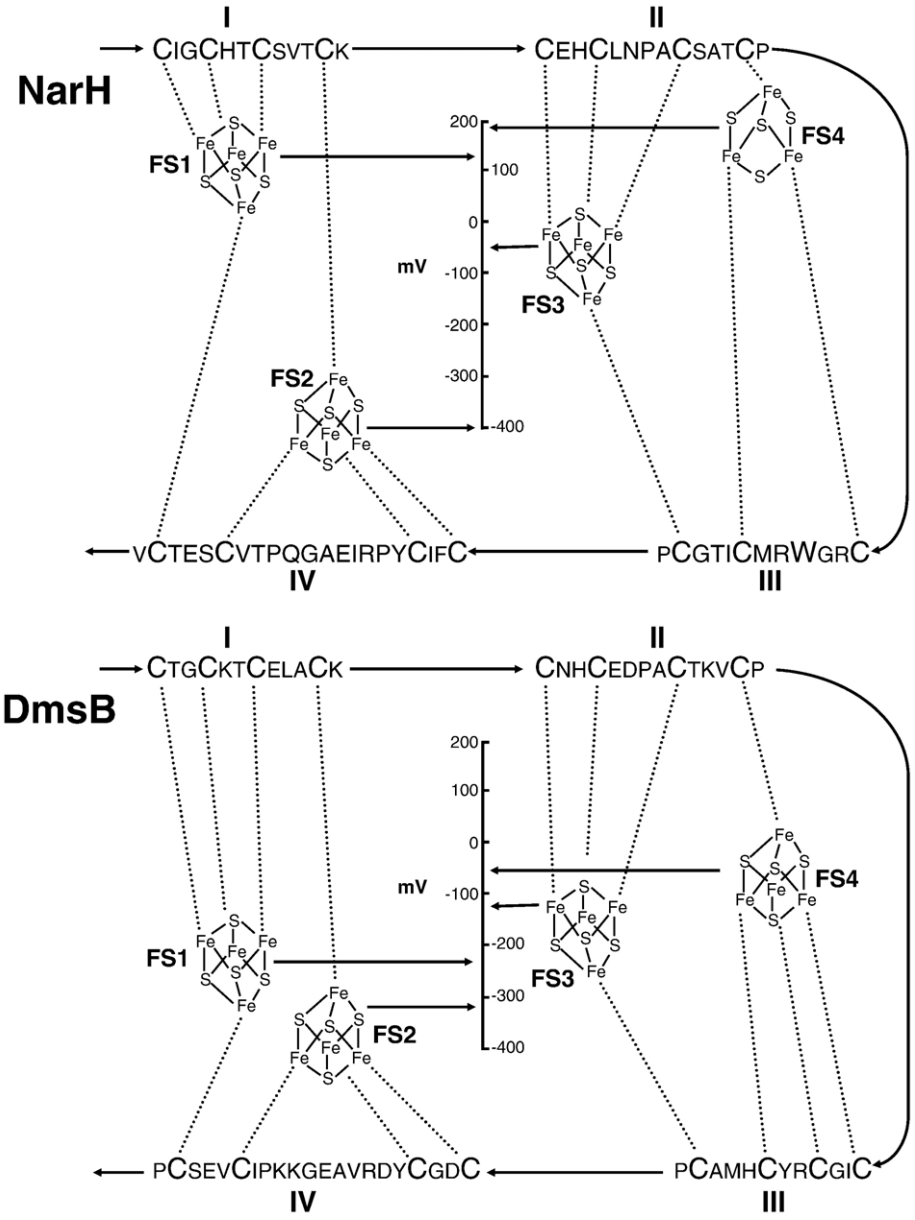

Fig. 6. Coordination of the potentiometrically identified [Fe-S] clusters by NarH (top) and DmsB (bottom). In both DmsB and NarH, the sequence of [Fe-S] clusters, from FS1 to FS4 is from left to right, whereas catalytic electron transfer occurs from right to left.

this context suggests the presence of an additional cluster (FS0) in DmsA having an E_m of approximately −140 mV. A second line of evidence for the presence of an FS0 cluster in DmsA is based on the observation by EPR of a [3Fe-4S] cluster in this subunit when the second Cys (DmsA-C67) residue of the Cys group is mutated to either a Ser or Ala residue [41]. The [3Fe-4S] cluster of a DmsA-C67S mutant has an E_m of approximately +170 mV [41], which can be predicted empirically to be approximately 300 mV higher than that of a [4Fe-4S] cluster in a similar overall coordination environment [61,64], resulting in an estimated E_m for FS0 of approximately −130 mV, which is consistent with the assertion that FS0 is the species observed interacting with the Mo(V) of the Mo-*bis*PGD cofactor. It is likely that the FS0 cluster of DmsA is recalcitrant to EPR detection in the wild-type enzyme, possibly due to it having a high spin ($S=3/2$) ground state in its reduced form. Alternatively, its spectrum may simply be too broad for it to be detected under the EPR instrument conditions used to date.

EPR spectra with conventional features around $g=2$ have been recorded for FS0 clusters in close relatives of the CISM catalytic subunits, such as the periplasmic nitrate reductase (NapA) and the formate dehydrogenase H (FdhF) catalytic subunit of the formate hydrogen lyase complex [65,66]. The consensus sequence for the FS0 coordinating Cys group is $(C_A/H_A)x_{2-3}C_Bx_3C_Cx_{27-34}C_Dx(K/R)$ [41], and it is notable that examples of FS0 with a low spin ground state have a Cys at the C_A position and two residues between C_A and C_B, whereas the one example of FS0 that has been demonstrated to have a high spin ground state has a His at the C_A position followed by three residues between this His and C_B [43]. In the case of

DmsA where FS0 has not yet been observed by EPR in its [4Fe-4S] cluster form, the C_A Cys residue is retained, but the spacing between C_A and C_B is, as in the case of NarG, three residues rather than two. It is also notable that a CISM-related enzyme, dimethylsulfide dehydrogenase (DdhABC) from *Rhodovulum sulfidophilum* also appears to have an FS0 cluster with a coordinating Cys group similar to that of NarG (with a His at the C_A position) and this cluster may also have a high-spin ground state [67,68]. At the present time there is a general paucity of available EPR data on FS0 clusters in CISM-type catalytic subunits other than those mentioned briefly above.

One question of fundamental importance surrounding electron transfer through the CISM type enzymes to or from the catalytic subunits is the identification of the precise pathway of electron transfer. Dutton and co-workers have developed theoretical models that predict intercenter electron transfer rates [69–71]. A couple of aspects of these models are intriguing. (i) The effective edge-to-edge distance limit for kinetically competent electron transfer appears to be approximately 14 Å. (ii) Components of the electron transfer relay (ETR) with unusually low E_m values do not necessarily impose a rate-limiting step upon overall catalysis. Active site catalysis involving obligatory chemical rearrangements in almost every case imposes the upper limit on enzyme turnover rate. (iii) The model for intercenter electron transfer treats the intercenter milieu as a conductor with a generalized "packing density". Given the complexities of protein structure this latter assumption appears reasonable in terms of simplifying the model. However, other studies suggest that specific bond-orbital electron transfer routes may exist between ETR components and that these have been selected for during evolution [72]. In the case of the CISM catalytic subunits, two studies have provided data that can be interpreted to support the specific electron transfer pathway model. In the case of DmsABC, mutagenesis of a highly conserved Arg residue that is predicted to be located between FS0 and PGD-P results in an enzyme that is unable to support growth on DMSO and has minimal menaquinol:DMSO oxidoreductase activity [40]. Interestingly, in this mutant enzyme (a DmsA-R61S mutant) the spin–spin interaction between the paramagnetic Mo(V) of the Mo-*bis*PGD cofactor and FS0 is eliminated [62], consistent with the pathway of this interaction being similar to that of electron transfer between the two centers. In the periplasmic nitrate reductase (NapA) from *Ralstonia eutropha*, mutagenesis of the residue equivalent to DmsA-R61, NapA-K56, also results in a dramatic decrease in catalytic turnover [73]. Although the total extent of the mutant effects on both DmsABC and NapA – secondary effects on the FS0 and Mo E_m values for example – is not currently known, it appears likely that the pathway of electron transfer between FS0 and the Mo-*bis*PGD cofactor does depend on the presence of specific amino acid residues. We anticipate that these issues will soon be resolved using a combination of site-directed mutagenesis, redox potentiometry, enzymology, and protein crystallography.

An important additional feature of the catalytic subunits of the CISM archetypes is the relationship between the presence of a twin arginine translocase leader (*tat* leader) at the N-terminus of the immature catalytic subunit and the final subcellular location of the mature catalytic dimer with respect to the cytoplasmic membrane [10–12]. When present, the leader sequence targets the catalytic and FCP subunits of the CISM archetypes to the periplasmic compartment, where they associate with their respective MAP subunits on the outer surface of the cytoplasmic membrane. Recently, it has been proposed that the N-terminus of NarG retains a sequence motif that resembles that of a *tat* leader [74]. This vestigial leader is uncleaved in the mature NarGHI complex in which the NarGH dimer is attached to the cytoplasmic side of the membrane-intrinsic NarI subunit. Although the specifics of holoenzyme maturation and cofactor insertion are outside the scope of this review, it is notable that the NarGHI system-specific chaperone (NarJ) has been shown to bind to the vestigial *tat* leader of NarG [75].

2.2. Structural diversity of the FCP subunits

The FCP subunit provides a molecular scaffold for four electron-transferring [Fe-S] clusters that shuttle electrons between the catalytic and membrane anchor subunits of an archetypal CISM and form the basis of the ETR connecting the two active sites of these enzymes. As mentioned in Section 1, these clusters are known as FS1–FS4 in sequence of increasing distance from the FS0 cluster of the catalytic subunit. Significant variation exists in the size of the CISM FCP subunit, ranging from 191 residues for the PsrB subunit of *Wolinella succinogenes* polysulfide reductase (PsrABC) [76] to 512 residues for the NarH subunit of *E. coli* nitrate reductase A [77]. Of the known CISM FCP subunits, only two have been subjected to scrutiny by redox potentiometry and EPR spectroscopy: the DmsB subunit from *E. coli* DmsABC [57–59,61] and the NarH subunit of *E. coli* nitrate reductase A [44–46,48,52,54]. The four [Fe-S] clusters of the FCP subunit are coordinated by four Cys groups (I–IV) with sequences similar to those found in the bacterial eight-iron ferredoxins ($C_AxxC_Bx_{2-11}C_Cx_3C_DP$) [78,79]. Coordination by the Cys groups occurs in such a way that iron atoms from each [4Fe-4S] cluster interact with three Cys residues from one Cys group (C_A–C_C) and a fourth from a second Cys group (C_D) [80,81]. In the case of NarH, one of the [Fe-S] clusters, FS4, is a [3Fe-4S] cluster, and is thus coordinated by two Cys residues from one Cys group (equivalent to C_A and C_C) and a third (equivalent to C_D) from another. Careful analyses of redox titration data from wild-type and mutant variants of NarH and DmsB have resulted in assignments of the potentiometrically identified clusters to the Cys groups as shown in Fig. 6. In the case of NarH, the coordination model has been confirmed by X-ray crystallography [7,9]. The coordination environments of the four [4Fe-4S] clusters of the FCP subunit of FdnGHI have also been determined by X-ray crystallography [8], but in this case there is an almost complete dearth of spectroscopic and biochemical data to complement the high-resolution structural data.

Inspection of the E_m values of the [Fe-S] clusters of NarH and DmsB reveals energy profiles through which electrons must pass on their journey from the MAP subunit to the FS0 cluster of the catalytic subunit [58]. Each cluster in the sequence from

FS4 to FS1 (and FS0) must undergo formal reduction during electron transfer. Simplistically, the energy barrier presented by a cluster increases as the ΔE_m between it and the preceding cluster in the chain becomes more negative [69–71]. It is notable that in both NarH and DmsB, the progress of E_m values through the FCP segment of the ETR does not follow the expected trend from lower to higher potentials. In both subunits, FS2 has a lower E_m value (−400 mV and −330 mV, respectively) than all the other components of the ETR, and thus represents the greatest thermodynamic barrier to catalytic electron transfer. The presence of a cluster of unusually low E_m value has also been noted in the ETRs of other *E. coli* respiratory enzymes, such as succinate:ubiquinone oxidoreductase (SQR, complex II) and menaquinol:fumarate oxidoreductase (QFR) [82–86]. It is possible that the thermodynamic barriers presented by such low-potential clusters may play a role in regulating catalytic turnover under conditions of low thermodynamic driving force. As stated above in Section 2.1, under standard *in vitro* conditions (i.e., high substrate concentrations), the thermodynamic barrier is unlikely to have a significant influence on electron transfer rates through the ETR.

Fig. 7A shows the results of an SSM comparison of the two currently available CISM FCP structures, NarH and FdnH. As is the case for SSM comparisons of NarG and FdnG, NarH is considerably larger than FdnH (509 versus 289 residues resolved in the respective structures) and contains a core structure of three segments of sequence that structurally align with an rmsd of 1.8 Å over 157 C-α atoms [9] (Fig. 7B). These three segments comprise two domains that coordinate the [Fe-S] cluster pairs FS1–FS2 and FS3-FS4, respectively. The additional three sequence segments that are absent from FdnH comprise three subdomains that are localized on the surface of the NarGHI dimer of heterotrimers. The role of these surface-localized domains in NarH remains unresolved, but as suggested in Section 2.1, they may serve to protect the ETR from damage caused by reactive nitrogen species. Finally, it should

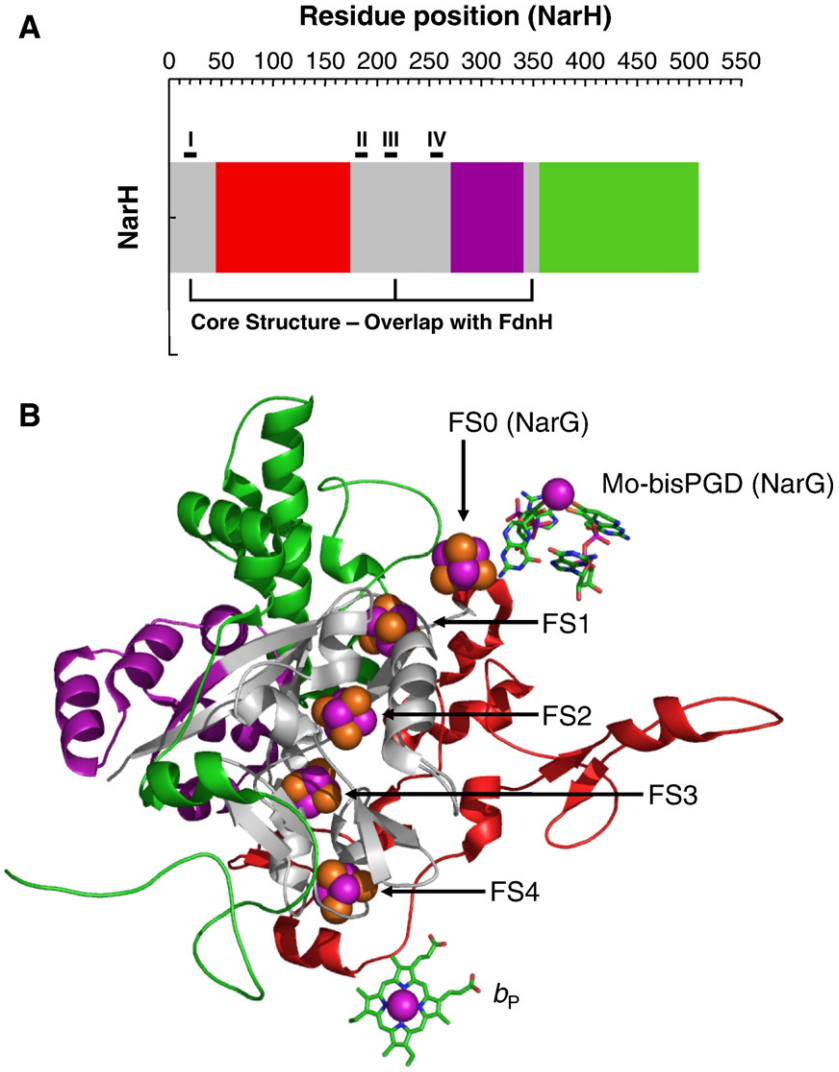

Fig. 7. Structural comparison of the NarH and FdnH FCP structures. (A) Graphical representation of the structural overlap between NarH (PDB file 1Q16) and FdnH (PDB file 1KQF). As indicated three sequence segments overlap structurally between the two proteins. (B) The location of the "core" FCP structure within NarH. The "core structure" is shaded grey. Color coding is as for panel A. For details of the alignment, see the text.

also be noted that FdnH also has additional sequence that is not part of the FCP core: there is a segment of sequence towards its C-terminus that forms a transmembrane (TM) segment in the holoenzyme structure [8,87].

2.3. Structural diversity of the MAP subunits

The MAP subunits exhibit the largest sequence and inferred structural diversity of the three subunits of the archetypal CISM. They range in size from approximately 217 residues for FdnI (Table 1) to 344 residues for the anchor (TtrC) of *Bordetella parapertussis* tetrathionate reductase (TtrABC). Biochemical and bioinformatic analyses suggest that these subunits may contain two hemes or none. In contrast to what is observed in the MAP subunits of the succinate:ubiquinone oxidoreductases [88–91], no CISM MAP subunits are currently known that contain only a single heme. Overall, the presence of heme correlates with the ability of the individual archetypes to participate in redox-linked proton translocation as described in Section 1. In the heme-containing examples, the Q-site is located distal to the side of the membrane that binds the catalytic and FCP subunits. In these cases, the function of the hemes is to provide a route for transmembrane electron transfer. The MAP subunits in general exhibit much poorer sequence motif retention than that exhibited by the catalytic and FCP subunits. Despite this, by using archetype sequences as bait for BLASTP searches [92], a large number of CISM MAP sequences can be identified and assigned to enzyme systems with a range of overall architectures.

2.3.1. NarI

Protein crystallography has revealed that *E. coli* NarI coordinates two hemes and has a transmembrane topology comprising five TM segments [9,93] (Fig. 8). One heme is located towards the cytoplasmic side of NarI and is referred to as the proximal heme (heme b_P), whereas the other is located towards the periplasmic side of NarI and is referred to as the distal heme (heme b_D) [54,94]. The two hemes function as a conduit for electron flow from a Q-site on the distal (periplasmic) side of NarI to the FS4 [3Fe-4S] cluster of NarH which is located on the cytoplasmic side of the membrane [49,51,53,54,56,93,94]. The structure of the protein reveals that the Q-site is located between transmembrane segments 2 and 3 towards the periplasmic side of NarI, and that the hemes are coordinated by four His residues, two of which are located in TM segment 2 and two in TM segment 5. This architecture results in the protons derived from quinol oxidation being released on the periplasmic side of the cytoplasmic membrane. Electrons derived from quinol oxidation at the Q-site are used to reduce nitrate at the Mo-*bis*PGD active site in a reaction that consumes two cytoplasmic protons. Thus the overall reaction catalyzed is:

$$QH_2 + 2H^+_{cytoplasmic} + NO_3^- \rightarrow Q + 2H^+_{periplasmic} + NO_2^- + H_2O$$

Thus, catalytic turnover results in the net translocation of one proton per electron from the cytoplasm to the periplasm.

BLASTP [92] searches against the SWISSPROT and TREMBL databases reveal ~90 unique sequences showing significant sequence similarity to NarI in ClustalW [95] alignments. These sequences can be assigned to a distinct family of hydrophobic diheme cytochrome *b*. Completely conserved residues within this family are indicated in Fig. 8. Although no significant sequence similarity is detected between NarI and the MAP subunits of the other archetypes of the CISM family, significant similarity is detected with the MAP subunit (HdrE) of a subclass of the HdrED-type heterodisulfide reductases [96]. In *Methanosarcina barkeri*, HdrE anchors a catalytic subunit (HdrD) that contains an eight-iron ferredoxin motif to the inner surface of the cytoplasmic membrane [97]. Analysis of ~27 HdrE homologs within the overall NarI family using the TMHMM algorithm (available at http://www.cbs.dtu.dk/services/TMHMM/) [98,99] reveals the presence of the 5 TM segment core of NarI, as well as an additional N-terminal TM segment.

2.3.2. FdnI/PhsC/CybH

As is the case with NarI, FdnI coordinates two hemes, but the architecture of its core structure is distinct from that of NarI, with the proximal heme (b_P) being coordinated by His residues in TM segment 2 and TM segment 4 (Fig. 8), and the distal heme (b_D) being coordinated by residues from TM segment 1 and TM segment 4 [8]. Because FdnG and FdnH are translocated across the cytoplasmic membrane by the *tat* translocon [100], they are located on the opposite side of the membrane compared to the NarG and NarH subunits of the NarGHI complex. This arrangement results in heme b_P being located towards the periplasmic side of the membrane and heme b_D being located towards the cytoplasmic side of the membrane (i.e., the opposite of what is observed in the NarGHI complex). The Q-site is sandwiched between TM segments 3 and 4 in the vicinity of the heme b_D towards the cytoplasmic side of FdnI (distal to the FdnGH catalytic dimer). Proton uptake thus occurs from the cytoplasm during formate-dependent quinone reduction, and this results in enzyme turnover contributing to the proton electrochemical potential across the cytoplasmic membrane [5] (Fig. 2). As indicated in Section 1, the availability of atomic-resolution structures of FdnGHI [8] and NarGHI [7,9] provides validation of the chemiosmotic redox loop mechanism for the generation of a transmembrane proton electrochemical potential [13]. The overall reaction catalyzed by FdnGHI can be summarized as follows:

$$HCO_2^- + Q + 2H^+_{cytoplasmic} \rightarrow CO_2 + H^+_{periplasmic} + QH_2$$

Sequence database searches using the *E. coli* FdnI sequence as bait return >140 unique hits. These include examples of the MAP subunit (PhsC) of *Salmonella typhimurium* thiosulfate reductase (PhsABC), which is a CISM archetype, and the CybH subunit of *E. coli* hydrogenase-1 (MbhSL-CybH), which is not. The core structure of the FdnI family of proteins is supplemented in many holoenzyme structures by a C-terminal TM segment of the FCP subunit that is part of the largely membrane-extrinsic catalytic dimer. In the case of the FdnGHI complex, a fifth TM segment is contributed by FdnH [5,87]. An almost identical situation exists in the case of the hydrogenase-1 class of enzymes, such as that of *W. succinogenes*, with the electron

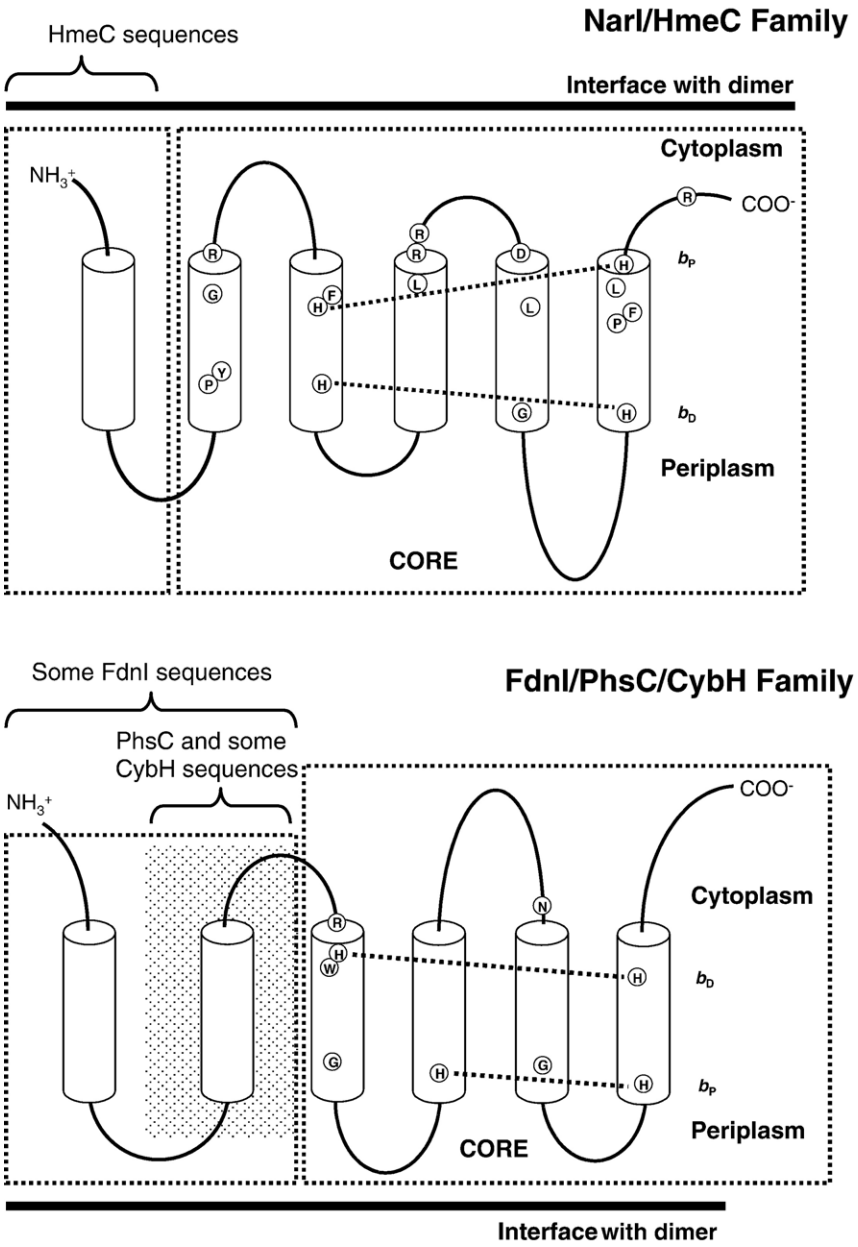

Fig. 8. Transmembrane topologies of the NarI/HmeC and FdnI/PhsC/CybH families of MAP subunits. 90 Unique sequences belonging to the NarI/HmeC family were mined from the SWISSPROT/TREMBL database using the sequence of *Escherichia coli* NarI as bait. 28 Sequences belong to the HmeC subgroup, of which 18 are predicted to have 6 rather than 5 TM segments. 140 Unique sequences belonging to the FdnI/PhsC/CybH family were identified using *E. coli* FdnI and *Wolinella succinogenes* PhsC as bait. 54 Sequences of these can be categorized as members of the FdnI subgroup, 14 as members the PhsC subunit, and 67 as members of the CybH subgroup. The approximate positions of residues are as indicated.

transfer subunit (MbhS) contributing a TM segment to a membrane-intrinsic domain that is essentially identical in architecture to that of FdnGHI [101,102]. In this case, although the electron transfer subunit does not belong to the FCP family, it does provide an electron transfer conduit from a nickel containing catalytic subunit that contains the site of hydrogen oxidation [103].

The MAP subunit of *S. typhimurium* thiosulfate reductase (PhsC) also belongs to the FdnI family. This enzyme is another heterotrimeric CISM archetype that is responsible for the reduction of a single thiosulfate to sulfide and a sulfite [104]. However, in this case the anchor is predicted to comprise 5 TM segments, all of which are part of PhsC (Fig. 8). The FCP subunit, PhsB, does not have a segment of hydrophobic amino acid residues sufficiently long to cross the cytoplasmic membrane. It has been proposed that the fifth TM segment of the membrane-intrinsic domain may be evolutionarily derived from the C-terminal TM segment of the FCP subunit (or the electron transfer subunit, MbhS, of the hydrogenase-1 family) [102].

A final variant of the FdnI-type architecture is also worth noting. Database mining reveals a number of MAP sequences assigned to FdnGHI type formate dehydrogenases in which the FCP lacks a C-terminal TM segment, and the MAP subunit is

predicted to have *six* TM segments. We were able to identify ~18 unique sequences corresponding to this architectural subset, including formate dehydrogenase MAP subunits from *Vibrio parahaemolyticus* and *Alkalilimnicola ehrlichei*.

On the basis of the above, the minimum functional membrane-intrinsic heme-coordinating domain containing the FdnI architecture appears to be a heme-coordinating 5 TM segment bundle. In the cases where additional TM segments are present, their function is currently unknown.

2.3.3. DmsC/NrfD/PsrC

BLASTP searches using the CISM MAP subunits of *E. coli* DMSO reductase (DmsABC) and *W. succinogenes* polysulfide reductase (PsrABC) as bait yield a set of 91 unique sequences that appear to share an identical overall architecture (Fig. 9). Each subunit is predicted to comprise 8 TM segments, and in the case of DmsC, this prediction is backed up by experimental results from β-lactamase (*blaM*) and alkaline phosphatase (*phoA*) gene fusions [105,106]. In addition to the two CISM archetypes, members of the NrfD subunit of the NrfABCD nitrite reductase system were identified as being members of the same overall family as DmsC and PsrC. In this case, however, NrfD forms a complex with an FCP subunit (NrfC) that does not form a catalytic dimer with a CISM catalytic subunit. Instead, NrfC interacts with a dimeric pentaheme cytochrome *c* subunit (NrfB) which shuttles electrons to a second dimeric pentaheme cytochrome *c* (NrfA) that contains the site of nitrite reduction [107,108].

With one important exception, bioinformatic and biochemical data suggest that the DmsC/NrfD/PsrC family of MAP subunits interact with partner subunits on the periplasmic side of the cytoplasmic membrane. In many cases, this is based on the presence of a *tat* leader sequence at the N-terminus of one of the partner proteins which directs the fully folded prosthetic group containing protein to the *tat* translocon [10,11,74]. However, controversy still surrounds the location of the catalytic dimer (DmsAB) of DmsABC, with a significant body of evidence suggesting that DmsAB is located at the *cytoplasmic* surface of the cytoplasmic membrane [60,106]. Contradictory data have been obtained by expressing histidine-tagged DmsAB in a Δ*dmsC* mutant [109]. In this case, DmsAB is detected in the periplasm by immunoblotting with an anti-histidine tag antibody, implying that it is attached to the periplasmic side of the membrane in the DmsABC holoenzyme. When a similar experiment is performed with non-tagged DmsAB, it remains cytoplasmically localized [110]. Bioinformatic analyses can also provide insights into the overall topology of DmsABC. Analysis of sequence conservation within the DmsC family of

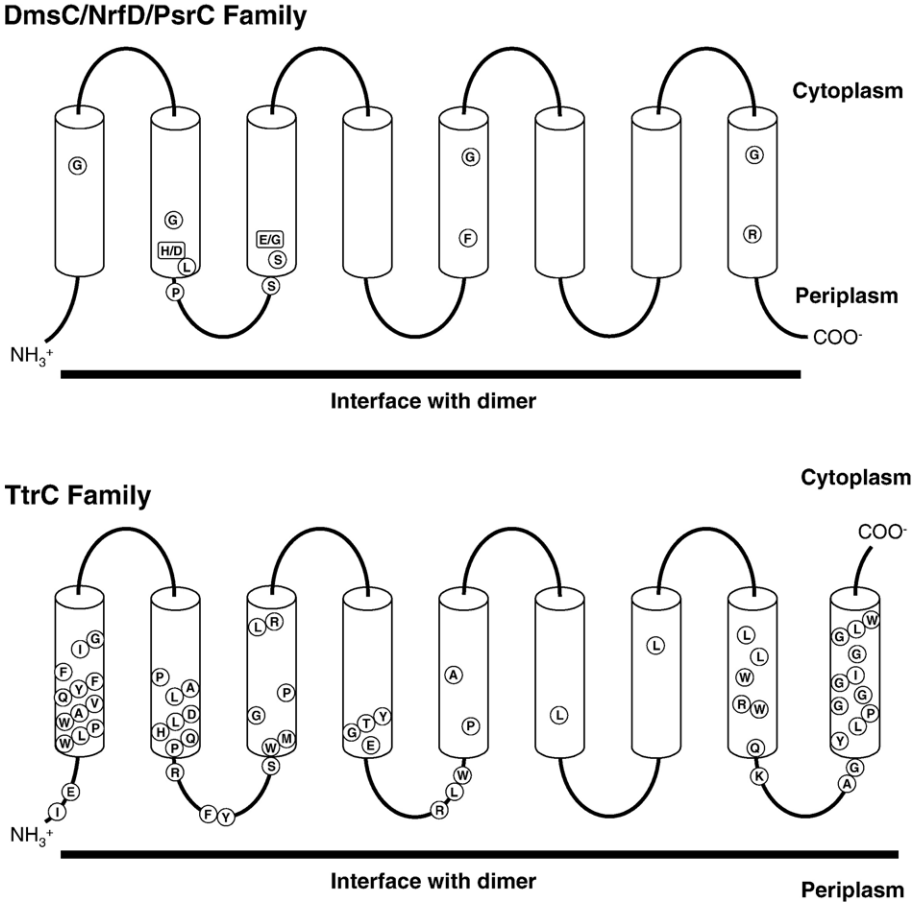

Fig. 9. Transmembrane topologies of the DmsC/NrfD/PsrC and TtrC families of MAP subunits. 91 Unique sequences belonging to the DmsC/NrfD/PsrC family were mined from the SWISSPROT/TREMBL database using the sequence of *Escherichia coli* DmsC as bait. Of these, 43 can be assigned as DmsC sequences, 28 as PsrC sequences, and 20 as NrfD sequences. 10 Sequences were identified as members of the TtrC family. The approximate positions of residues are as indicated.

proteins, in comparison with the MAP subunits of range of well characterized bacterial redox enzymes, reveals that conserved residues that may define subunit–subunit interactions are biased towards periplasmically localized loops between TM helices [111]. This implies a periplasmic location for the DmsAB dimer.

In the proposed topology of DmsC [105], amino acid residues proposed to be involved in menaquinol binding and oxidation are localized towards the periplasmic side of the subunit. These include DmsC-H65 that has a proposed location towards the periplasmic end of TM segment 2 [61,112] and DmsC-E87 that has a proposed location towards the periplasmic end of TM segment 3 [113]. Given that there is no evidence for the presence of hemes in DmsC that might be able to form an electron-transferring conduit from a periplasmically oriented Q-site to the [Fe-S] clusters of the FCP subunit (DmsB), the simplest explanation for the apparently conflicting sets of data is that the DmsAB catalytic dimer is indeed localized on the periplasmic face of the cytoplasmic membrane. This conclusion is all the more compelling when the DmsABC system is compared to the NrfABCD system, in which the DmsC-like MAP (NrfD) clearly interacts with a periplasmically oriented DmsB-like FCP subunit (NrfC) [27].

2.3.4. TtrC

Tetrathionate reductase (TtrABC) has a MAP subunit that is predicted to have 9 TM segments. BLASTP searches using the sequence of TtrC as bait return a relatively small number of 10 unique sequence hits in the SWISSPROT and TREMBL databases. Sequence similarity within this group appears to be clustered towards the periplasmic side of the cytoplasmic membrane, consistent with the catalytic TtrAB dimer being directed by the *tat* leader sequence on TtrA to the periplasmic compartment (Fig. 9). Given the relative lack of biochemical data on TtrABC systems [114], it is not clear if any of the sequences identified as members of the TtrC family are involved in other redox systems.

Comparison of the conserved residues indicated in the DmsC/NrfD/PsrC family with those of the TtrC family reveals that the two may be related. However, a number of conserved residues within the TtrC family are not conserved in the DmsC/NrfD/PsrC family and vice versa (Fig. 9), and there are no conserved heme-coordinating His residues to provide additional evidence of relatedness. Overall, it is possible that the architecture of TM1-TM8 in both cases may be essentially identical.

3. Variations on the CISM theme

One remarkable aspect of prokaryotic electron transfer chain enzymes is the way individual subunits evolved to take on specific functions such as soluble substrate redox reactions, electron transfer, quinone redox chemistry, and bioenergetically specific quinone-dependent proton deposition and uptake. As stated above, in the CISM family of enzymes, these functions are carried out by the catalytic, FCP and MAP subunits, respectively. However, evolution has resulted in considerable mixing and matching of these subunits throughout a range of redox enzymes. A consequence of this is that structural modules of the CISM archetypes appear in a diverse range of enzymes [107,115] (Fig. 10; Table 2B). The catalytic subunit architecture appears in soluble oxidoreductases such as *E. coli* TMAO reductase (TorA) [116,117] and *Rhodobacter* DMSO reductase (DorA) [14,15], or they can pair up with electron transfer subunits containing hemes (*E. coli* periplasmic nitrate reductase, NapAB [33,65,118]) or Rieske-type [2Fe-2S] clusters (*Alcaligenes faecalis* arsenite oxidase (AoxAB) [119]). In the case of the tungsten-containing formate dehydrogenase from *Desulfovibrio gigas* (W-FdhAB), the catalytic subunit pairs up with a subunit similar in structure to that of the archetypal FCP, except that it contains only three [4Fe-4S] clusters and is referred to herein as a three-cluster protein (TCP) [120]. In the ethylbenzene dehydrogenase from *Aromatoleum aromaticum*, a catalytic subunit (EbdA) and an FCP subunit (EbdB) form a complex with a cytochrome *b* subunit (EbdC) [39]. The FCP subunit is conserved in additional respiratory systems that do not contain a Mo-*bis*PGD binding subunit. For example, in the NrfABCD nitrite reducing system of *E. coli*, the FCP subunit (NrfC) appears to be attached to a membrane anchor (NrfD), but electron transfer to a soluble nitrate-reducing cytochrome *c* (NrfA) is mediated by a cytochrome *c* subunit (NrfB) that appears to be part of the membrane-associated complex of this system [26,107]. In another variation on this theme, in a *Desulfovibrio vulgaris* hydrogenase system, an FCP subunit interacts with a high molecular weight cytochrome *c* (Hmwc) [121–123], which is likely to interact with a soluble periplasmic hydrogenase (comprising PhfL and PhfS) [107]. Thus, the FCP subunit is of particular importance due to its occurrence in multiple respiratory chains in a range of bacteria. Two final cases of subunit modularity involve large multisubunit (>3) assemblies. The first is the formate hydrogen lyase of *E. coli* (FHL), in which the catalytic subunit interacts with an FCP (HycB), which itself interacts with multiple components of the *E. coli* hydrogenase-3 encoded by the *hycABCDEFGHI* operon. The second is the unique case of the NuoG/Nqo3 subunit of prokaryotic and mitochondrial complex I, in which the overall structure of the catalytic subunit is retained as a component of the membrane-extrinsic arm of the complex, but where the Mo-*bis*PGD cofactor itself has been lost during evolution [124,125]. Selected examples of these variations are described in more detail below.

3.1. Enzymes that contain a CISM catalytic dimer

A number of bacterial enzymes exist that contain one or more subunits of the archetypal CISM architecture (Table 2B, Fig. 10). An important subclass of these includes those enzymes that contain the catalytic and FCP subunits. These have an overall architecture comprising only the CISM catalytic dimer, or they may comprise a CISM catalytic dimer complexed with a non-membrane integral cytochrome *b* subunit. An example of the former case is the periplasmic tungsten-containing formate dehydrogenase (W-FdhAB) of *D. gigas* [120]. An example of the latter is found in the *Rhodovulum sulfidovilum* dimethylsulfide dehydrogenase (DdhABC) [67,68].

Sequence and biophysical analyses suggest striking similarities between the core architecture of the CISM catalytic dimer

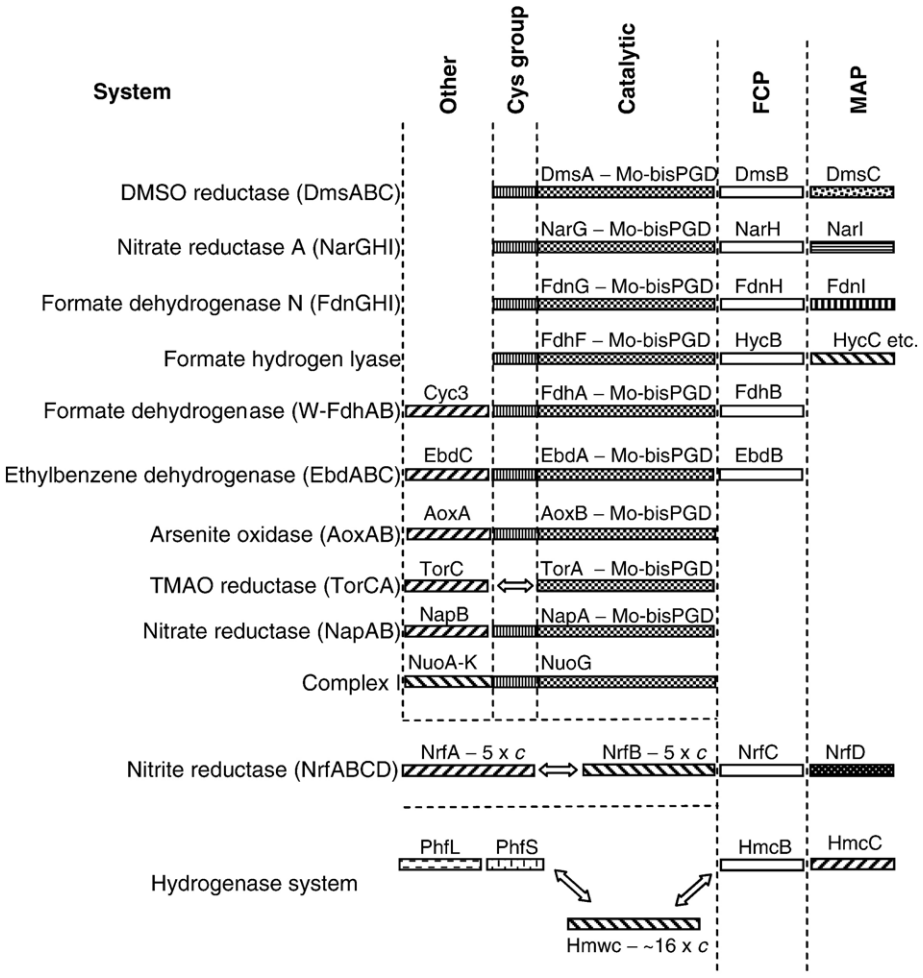

Fig. 10. Modularity of the CISM core subunits. The interactions between the subunits of the indicated systems are indicated, with the common subunit architectures being separated by the dashed lines. ⇔ Implies a transient interaction between the indicated subunits, possibly mediated by other proteins such as periplasmic *c*-type cytochromes.

subunits of DdhABC (DdhAB) and that of the archetypal heterotrimeric CISM enzymes [67]. DdhAB contains a Mo-*bis*PGD cofactor along with a complement of 5 predicted [Fe-S] clusters, with FS0-FS3 being [4Fe-4S] clusters and FS4 being a [3Fe-4S] cluster. The presence of a *tat* leader sequence directs the DdhAB subunits to the periplasmic compartment, where they associate with a soluble monoheme cytochrome *b* to form a functional heterotrimer. Sequence analyses suggest that the latter subunit appears to be directed to the periplasmic compartment by the *sec* system. It is notable that in this case both the *tat* and *sec* systems appear to be necessary for the translocation and assembly of mature DdhABC. *Thauera selenatis* selenate reductase (SerABC), *Ideonella dechloratans* chlorate reductase (ClrABC), and *Azoarcus* sp. *EB1* ethylbenzene dehydrogenase (EbdABC) appear to have a similar subunit composition to that of DdhABC. The structure of the EbdABC complex from *A. aromaticum* has recently been reported, confirming previous predictions of the cofactor compositions of this group of enzymes [39]. In the case of EbdABC, no *sec* leader is predicted for the EbdC subunit, suggesting that in this case the entire trimer may be translocated by the *tat* system. This suggests an evolutionary convergence of subunit assembly to the cytoplasmic compartment, resulting in an increase in the number of "passenger proteins" that are directed to the periplasmic compartment by a single *tat* signal sequence attached to only one subunit of the multisubunit complex.

D. gigas formate dehydrogenase (W-FdhAB) is an enzyme comprising a CISM catalytic dimer with a W-*bis*PGD cofactor and FS0 cluster in the catalytic subunit (W-FdhA) paired with an FCP-related subunit (W-FdhB) that contains only three [Fe-S] clusters (FS1–FS3) [120]. It can thus be referred to as a TCP. The W-FdhAB catalytic dimer is directed to the periplasmic compartment by a *tat* leader located at the N-terminus of W-FdhA and is believed to donate electrons derived from formate oxidation to a periplasmic cytochrome *c* that co-purifies with the W-FdhAB dimer [126–128]. Comparison of the N-terminal sequence of the co-purified cytochrome *c* subunit [128] with the SWISSPROT/TREMBL databases reveals it to be a tetraheme cytochrome c_3 [129].

Another important variant of the CISM catalytic dimer architecture is presented by the pyrogallol–phloroglucinol transhydroxylase (PgtLS) of *Pelobacter acidigallici*. The interconversion of pyrogallol and phloroglucinol does not constitute a net redox reaction, but the reaction mechanism combines an

oxidative hydroxylation with reductive dehydroxylation via the Mo-*bis*PGD cofactor [130,131]. PgtLS comprises a modified CISM catalytic dimer in which a Mo-*bis*PGD cofactor is present in PgtL, along with FS1, FS2, and FS3 in PgtS, but FS0 and FS4 are absent [132]. Thus, as is the case with W-FdhAB, the FCP-related subunit is a TCP. Another interesting feature of PgtLS is that its TCP subunit contains a C-terminal domain with a fold similar to that of the cell-adhesion protein fibronectin III [133]. Based on this similarity, it has been suggested that PgtLS is membrane associated via this additional domain [130,132]. The vestigial nature of the FCP-related TCP subunit and the absence of FS0 from its catalytic subunit suggest that the PgtLS enzyme architecture arose via an evolutionary degradation of the cluster composition of a typical FS0–FS4 ETR.

It is notable that when the structure of NarH (1Q16, chain B) is used as bait to search the entire PDB (www.pdb.org) archive using the SSM server, the structure of the FCP-related subunit PgtS is the closest structural homolog of NarH in the PDB archive. The structure of NarH aligns to that of PgtS (1TI2, chain B) with an rmsd of 1.7 Å over 177 C-α atoms (out of a total of 274 for PgtS). The tungsten-containing formate dehydrogenase of *D. gigas* (1H0H, chain B) aligns with NarH with an rmsd of 1.9 Å over 156 C-α atoms (out of a total of 214 for W-FdhB). The bona fide FCP subunit FdnH aligns with an rmsd of 1.8 Å over 157 C-α atoms. Fig. 11 shows the structural alignments of NarH with PgtS and W-FdhB. It has been proposed that the structures of PgtS and W-FdhB arose evolutionarily from an archetypal FCP protein [132].

3.2. Enzymes containing a Mo-bisPGD catalytic subunit

Enzymes containing only the catalytic subunit of the CISM architecture include the periplasmic TMAO reductase (TorA) [116,117], the periplasmic nitrate reductase (NapAB) [38,134–137], the cytoplasmic biotin-*d*-sulfoxide reductase (BisC) [138,139], and the cytoplasmic arsenite oxidase (AoxAB) [119,140]. These enzymes have been reviewed extensively (see [16,140,141]), and their catalytic subunit structures very similar to those available for DorA [14,15]. The following points warrant mention in the context of this review. (i) NapB forms a heterodimer with NapA, but NapB is translocated to the periplasmic compartment via the *sec* translocon, whereas NapA is translocated via the *tat* translocon. Thus, as in the case of the dimethylsulfide dehydrogenase family, multiple components of the periplasmic nitrate reductase appear to cross the cytoplasmic membrane via the two primary translocons. (ii) The arsenite oxidase catalytic subunit (AoxB) forms a dimer with a Rieske type [2Fe-2S] cluster containing subunit (AoxA) that has the *tat* leader at its N-terminus. Thus, in this case, the Mo-*bis*PGD-containing subunit is carried across the cytoplasmic membrane

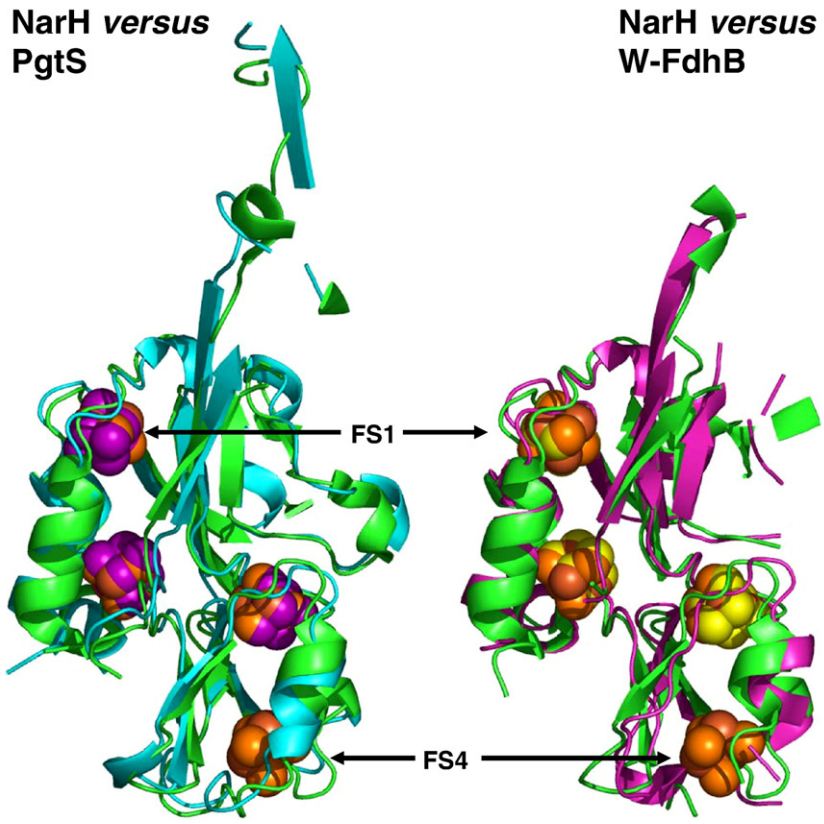

Fig. 11. Structural alignment between NarH (1Q16) and PgtS (1VLF), and between NarH (1Q16) and W-FdhB (1H0H). Note that in both PgtS and W-FdhB, the cluster corresponding to FS4 is missing. In both cases, NarH is colored green. PgtS is colored blue, and W-FdhB is colored purple. The positions of FS1 and FS4 are marked. In the case of PgtS, the [Fe-S] clusters are colored purple, and in the case of W-FdhB, the clusters are colored yellow. The NarH clusters are colored orange in both panels. For further details, see the text. Non-aligned structure is omitted for clarity. The structural alignment shown was created using the program PYMOL (DeLano Scientific LLC, San Francisco, CA).

by its partner in the arsenite oxidase complex. This contrasts with the vast majority of cases, where it is the Mo-*bis*PGD containing subunit that bears the *tat* leader and carries the subsidiary subunits across the membrane as "passenger proteins" [11,12,74].

3.3. Respiratory complex I (NADH:ubiquinone oxidoreductase) contains elements of the CISM catalytic subunit architecture

It has been noted that significant similarity exists between the C-terminal sequence of the Nqo3/NuoG subunit of bacterial and eukaryotic NADH:ubiquinone oxidoreductase (complex I) and the archetypal CISM catalytic subunit [125]. Additionally, the N-terminus of Nqo3/NuoG shares sequence similarity with the HoxU subunits of the iron-only hydrogenase from *Clostridium pasteurianum* [142] and the NAD^+-reducing [Ni-Fe] hydrogenase (H_2:NAD^+ oxidoreductase) of *R. eutropha* [143]. The recent determination of the structure of the membrane-extrinsic arm of *Thermus thermophilus* complex I enables direct structural comparisons to be made [124,144]. Table 2 shows the results of a SSM search of the PDB archive using the structure of Nqo3/NuoG as bait. The FdhF subunit of the *E. coli* formate hydrogen lyase (FHL) complex appears to be the highest scoring hit with an rmsd of 2.41 Å over 423 C-α atoms. Other high-scoring hits include NapA, AoxB, W-FdhA, PgtL, FdnG, and NarG. An obvious striking difference between the bait structure and the hits is the lack of a Mo-*bis*PGD cofactor in the former [124].

The Nqo3/NuoG subunit of *T. thermophilus* contains four [Fe-S] clusters (3 [4Fe-4S] clusters and 1 [2Fe-2S] cluster) [124,144], three of which are located in the N-terminal HoxU-like portion of the subunit. The fourth, N7, is only observed in a

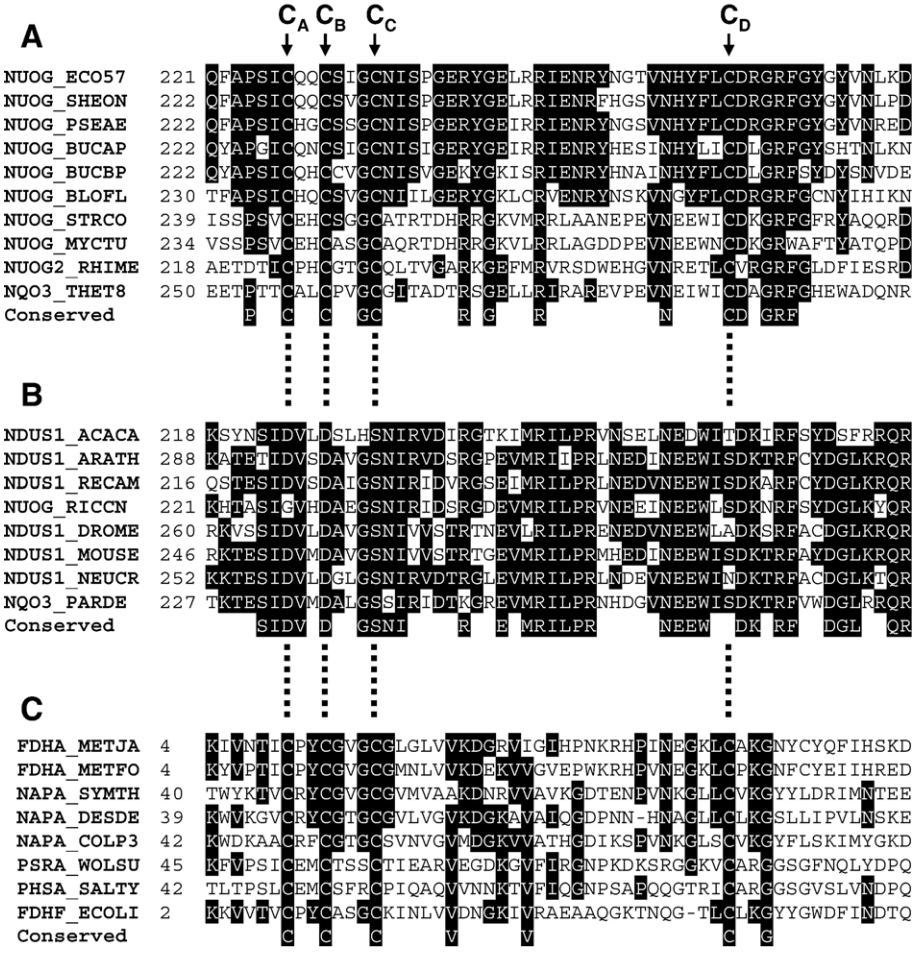

Fig. 12. Sequence alignment surrounding the [4Fe-4S] N7 cluster of complex I. The sequence of the Nqo3 subunit from *Thermus thermophilus* was used as bait to collect a range of Nqo3 subunit sequences from a broad phylogeny of organisms represented by annotated entries in the SWISSPROT database. The hits were filtered pairwise to remove sequences with ≥70% sequence identity to other hits. The resultant sequence set was aligned using ClustalX [95,146,147]. (A) Sequences from organisms containing a tetranuclear N7 cluster in their Nqo3/NuoG subunits. (B) Alignment of Nqo3/NuoG sequences that do not coordinate an N7 cluster. (C) Alignment of the FS0-coordinating sequences of CISM catalytic subunits that were returned using the *T. thermophilus* Nqo3 sequences as bait in the BLASTP search. The sequences aligned for each of the three panels were from *Escherichia coli O157:H7*, *Shewanella oneidensis*, *Pseudomonas aeruginosa*, *Buchnera aphidicola* subsp. *Schizaphis graminum*, *Buchnera aphidicola* subsp. *Baizongia pistaciae*, *Blochmannia floridanus*, *Streptomyces coelicolor*, *Mycobacterium tuberculosis*, *Rhizobium meliloti*, *Thermus thermophilus*, *Acanthamoeba castellanii*, *Arabidopsis thaliana*, *Reclinomonas americana*, *Rickettsia conorii*, *Drosophila melanogaster*, *Mus musculus*, *Neurospora crassa*, *Paracoccus denitrificans*, *Methanococcus jannaschii*, *Methanobacterium formicicum*, *Symbiobacterium thermophilum*, *Desulfovibrio desulfuricans*, *Colwellia psychrerythraea*, *Wolinella succinogenes*, *Salmonella typhimurium*, *E. coli* (in the order presented in the alignment). Vertical dashed lines indicate the positions of the C_A, C_B, C_C, and C_D residues. Conserved residues within each panel are highlighted.

Table 3
SSM search results using the structure of the complex I Nqo3 subunit as bait[a]

Hit: subunit and PDB code	System type	rmsd (Å)	N_{align}[b]
FdhF (*E. coli*) formate hydrogen lyase (1FDO)	Multisubunit complex	2.41	423
NapA (*D. desulfuricans*) periplasmic nitrate reductase (2NAP)	Monomeric reductase	2.64	428
AoxB (*A. faecalis*) arsenite oxidase (1G8K)	Dimer with Rieske type [Fe-S] protein (AoxA). FS0 is a [3Fe-4S] cluster	2.82	417
FdhA (*D. gigas*) W-containing formate dehydrogenase (1H0H)	Dimer with TCP (FdhB)	2.67	421
PgtL (*P. acidigallici*) pyrogallol–phloroglucinol transhydroxlase (PPTase) (1TI2)	Lacks FS0, dimer with TCP	3.85	376
FdnG (*E. coli*) FdnGHI (1KQF)	CISM archetype	2.64	430
NarG (*E. coli*) NarGHI (1Q16)	CISM archetype	3.18	416

[a] Coordinates contained in the PDB file 2FUG [124] were used as bait.
[b] Number of residues aligning with quoted rmsd, out of a total of 783 residues of the *Thermus thermophilus* Nqo3 subunit.

subset of bacterial species [145]. It is located approximately 21 Å from the nearest cluster of the Complex I ETR, which is the N5 [4Fe-4S] cluster, and does not appear to be on a plausible route of electron transfer through the complex I ETR [69]. Interestingly, the N7 cluster is located, as is the case with the bacterial CISM catalytic subunits, in close juxtaposition to the large internal cavity that would normally contain the Mo-*bis*PGD cofactor.

Fig. 12 illustrates a sequence alignment of a subset of sequences returned in a BLASTP search using that of Nqo3 as bait. As expected, the four cluster-coordinating Cys residues are conserved with spacings identical to those of the FS0-containing CISM subunits that were used in the alignment (generated using ClustalX [95,146,147]) (cf. panels A and C). In the CISM catalytic subunit family, there is a conserved Lys/Arg located two positions after that of the FS0-coordinating C_D residue. In both *E. coli* DmsABC and *R. eutropha* NapA, this residue has been demonstrated to be critical for electron transfer to the Mo-*bis*PGD cofactor [40,62,73]. Interestingly, this residue is completely conserved as a Lys in the group of Nqo3/NuoG sequences lacking the N7 cluster (Fig. 12B) and is conserved in 7 out of 10 sequences that contain the N7 cluster (Fig. 12A). However, in this group of sequences, there is an absolutely conserved Arg residue four residues after the C_D. The role of these conserved basic residues in the Nqo3/NuoG subunit family is currently unknown.

Comparison of the Nqo3/NuoG sequences that contain the N7 cluster with those that do not (cf. Fig. 12A and B) reveals that the C_A, C_B, C_C and C_D residues are changed to Asp, Asp, Ser, and Ser, respectively, when no cluster is present. The sequences lacking N7 also appear to have a large number of additional conserved basic and acidic residues in this region of the protein, with 7 conserved Asp/Glu residues and 6 conserved Arg/Lys residues, suggesting that this region of the protein may be stabilized by a large number of salt bridges and/or an unusually extensive hydrogen bonding network.

As noted above, the *T. thermophilus* Nqo3 subunit aligns structurally with a range of Mo-*bis*PGD containing catalytic subunits, including those of the CISM archetypes NarG and FdnG. These structural alignments are over 416 C-α atoms with an rmsd of 3.18 Å and over 430 C-α with an rmsd of 2.64 Å, respectively (Table 3). Fig. 13 illustrates graphical representations of these alignments and indicates the positions of the NarG and FdnG residues that are within 5 Å of the Mo-*bis*PGD cofactor or the FS0 [4Fe-4S] cluster. A strong correlation is observed between conservation of structure and the positions of residues involved in binding the two CISM catalytic subunit prosthetic groups. These observations support the assertion, based on sequence bioinformatics, that the Nqo3/NuoG architecture is evolutionarily derived from that of the CISM catalytic subunit [125]. It has been noted that the subunits of the hydrogenase components of the FHL complex encoded by the *hycA-I* operon share extensive sequence and inferred structural similarity with multiple complex I subunits [148]. Thus, the FHL complex comprising FdhF and the multiple subunits of hydrogenase-3 may be an important stepping stone in the evolution of mitochondrial complex I (see Section 6.1).

3.4. Systems containing FCP and MAP subunits

The *E. coli* periplasmic nitrite reductase system catalyzes the 6 electron reduction of nitrite to ammonium [149]. The nitrite reducing component of this system is a pentaheme cytochrome *c* (NrfA) in which four hemes are coordinated by the classic heme *c* binding motif (CXXCH), with the fifth ligand binding

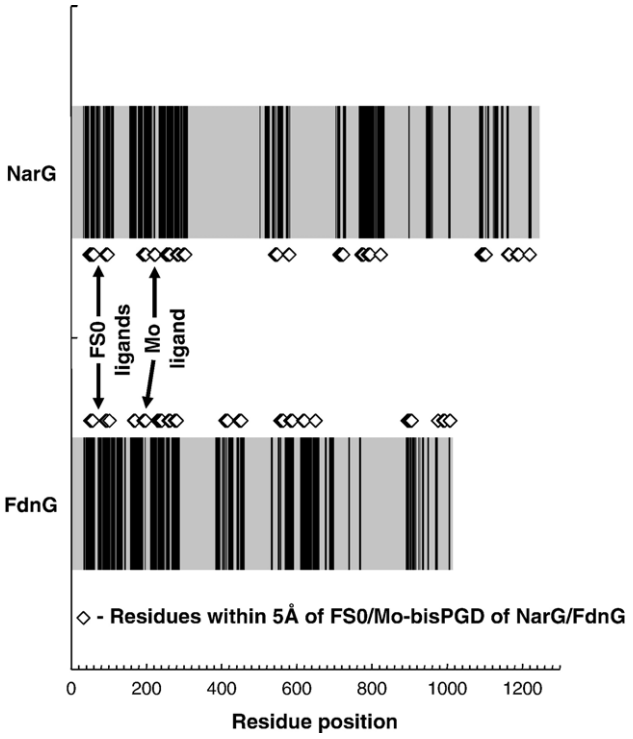

Fig. 13. Structural alignment of NarG (1Q16) and FdnG (1KQF) with Nqo3 (2FUG). Blocks of aligned structure are shown in black. Diamonds indicate the positions of residues within 5 Å of the Mo-*bis*PGD or FS0 of NarG and FdnG.

heme coordinated by a modified CXXHK motif [150,151]. This subunit accepts electrons from a membrane-bound complex containing a second pentaheme cytochrome *c* (NrfB), an FCP subunit (NrfC), and a MAP subunit (NrfD) belonging to the DmsC/PsrC/NrfD family. NrfC and NrfD appear to form a heterodimer, with the NrfC subunit providing an ETR through its four [4Fe-4S] clusters from a Q-site located in NrfD to a NrfB-binding site on the surface of NrfC [4,26,108]. NrfC is targeted to the periplasmic compartment by a *tat* leader.

E. coli hydrogenase-1 is a heterotrimer comprising a nickel-containing catalytic subunit (HyaB) that forms a membrane-extrinsic catalytic dimer with an electron transfer subunit (HyaA) that contains three [Fe-S] clusters (it contains two [4Fe-4S] clusters and a [3Fe-4S] cluster [152]). In the *W. succinogenes* form of the enzyme, the membrane-extrinsic dimer of HyaAB is clearly directed to the periplasmic compartment by a *tat* leader at the N-terminus of HyaA [153]. The hydrogenase-1 MAP subunit is a member of the FdnI/PhsC family, with two *b*-type hemes being coordinated by four His residues located in three TM segments [8,101,154]. Although HyaA is not an FCP, it shares the C-terminal TM segment of FdnH that forms part of the 5 TM segment membrane-intrinsic domain that is conserved between HyaABC and FdnGHI.

The final variant of the FCP-MAP architecture to be considered here is part of the hydrogenase system of *D. vulgaris* [121,122]. Its CISM-derived subunit comprises a *tat*-exported FCP (HmcB) anchored to the outside of the cytoplasmic membrane by a 10 TM segment MAP subunit (HmcC). The subunits appear to couple the quinone pool to a periplasmic Fe-hydrogenase comprising the PhfL and PhfS subunits via a soluble high molecular weight cytochrome *c* (Hmwc) that contains 16 hemes [155].

The FCP subunit architecture appears to be present in a relatively large number of bacterial electron transport systems ranging from the CISM archetypes themselves to systems where each of their redox partners is quite distinct from the archetypal catalytic and MAP subunits. However, for unknown reasons, their presence appears to be limited to prokaryotic organisms. In eukaryotes, with the exception of the very long [Fe-S] cluster-mediated ETR in mitochondrial complex I [124,144], the more typical ETR architecture is that exemplified by succinate: ubiquinone oxidoreductase (complex II), which comprises three [Fe-S] clusters [88,156,157].

4. Phylogenic relationships

Within the archetypal CISMs described above in Section 2 and listed in Table 2A, it is clear that each shares a very similar catalytic subunit and FCP subunit architecture. To obtain a greater understanding of the relationships between members of the catalytic and FCP families, we mined the SWISSPROT and TREMBL databases using the sequences of the CISM archetypes listed in Table 2A as bait (using the BLASTP search engine [92]). A total of 906 and 910 sequences were returned in these searches. To simplify further analyses by reducing the size of the dataset, these sequences were filtered by pairwise comparison to remove sequences with greater than 70% sequence identity to any other sequence in each set. Each sequence set was then aligned using the ClustalX program [95,146,147], and obvious outliers and truncates were removed. This approach yielded a catalytic subunit sequence set comprising 246 members and an FCP subunit set comprising 207 members. These two sequence sets were aligned and bootstrapped neighbor-joining trees were generated using the ClustalX program. Because the MAP subunits exhibit a far greater divergence of sequence and protein fold than the CISM catalytic and FCP subunits, we have focused our phylogenic analyses on the latter two families of proteins.

4.1. The catalytic subunits

Fig. 14 illustrates the phylogenic relationships between the catalytic subunit sequences obtained as described above. The filtered sequence set can be divided into 15 distinct clades. In the absence of available functional information, substrates were assigned to each member of each clade by searching each sequence against the SWISSPROT database (http://us.expasy.org/tools/blast/) using the BLASTP program [92]. The individual clades are summarized as follows.

4.1.1. NarG

The catalytic subunit of the CISM archetype NarGHI is typified by that found in *E. coli* and is described in more detail in Section 2.1. Members of the NarG family typically comprise approximately 1250 amino acids, rendering them considerably larger than the other CISM catalytic subunits. As described in Section 2.1, the structure of NarG contains a core similar to that found in DorA, and this is supplemented with surface-localized domains that both shield the ETR and provide a relatively narrow substrate binding funnel (Fig. 5) [7,9]. Additional characteristics include the presence of an FS0 cluster coordinated by an N-terminal Cys group that has until recently proven to be recalcitrant to EPR characterization [43,55]. The FS0-coordinating Cys group has a conserved His residue at the C_A position, and three residues are found between this His and the Cys at the C_B position. Additionally, as described in Section 2.1, the protein–Mo ligand is a conserved Asp residue. Finally, this group lacks an N-terminal *tat* leader. Overall, these characteristics allow facile identification of putative NarG sequences.

NarG sequences were identified in 24 species in the bacterial and archeal domains of life. Sequences were identified in two thermophilic archaea, *Pyrobaculum aerophilum* and *Aeropyrum pernix*. Of the bacterial phyla, 12 sequences were identified in phyla lacking an outer membrane (see Section 5 for discussion of bacterial morphology nomenclature), 8 from the Actinobacteria and 4 from the Firmicutes. Bacteria with outer membranes are represented by one Deinococcus-Thermus sequence (from *T. thermophilus*) and 9 Proteobacteria sequences. The broad distribution of NarG encoded by archeal and bacterial genomes is likely a reflection of the early evolutionary emergence of nitrate respiration.

4.1.2. DdhA/SerA/EbdA

Inspection of Fig. 14 reveals that these sequences represent a number of discrete branches along the branch leading to/from

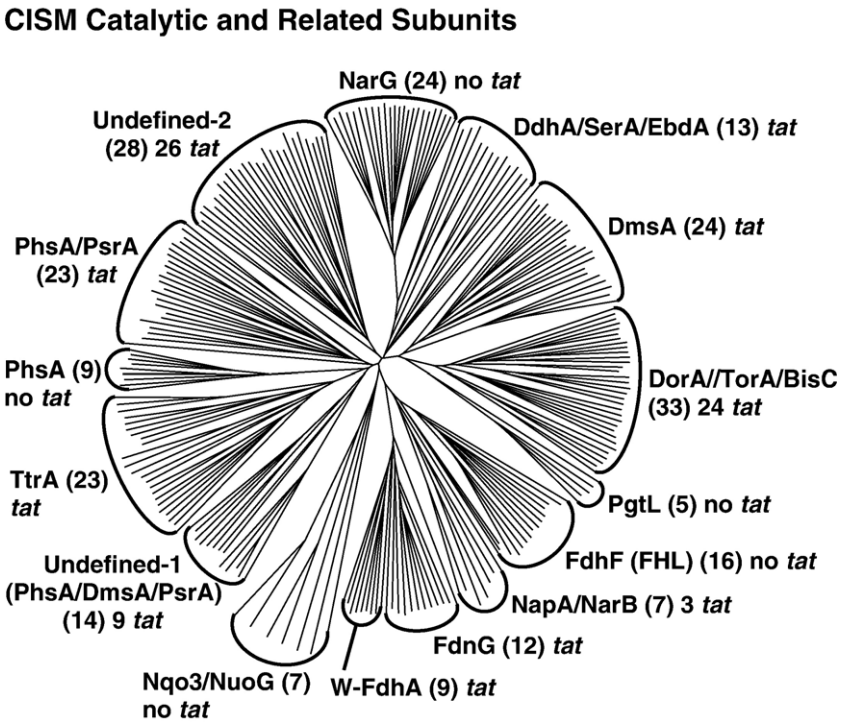

Fig. 14. Phylogenic analysis of the CISM catalytic and related subunits. Sequences were mined and analyzed as described in the text.

the NarG type subunit architecture rather than being a distinct clade in its own right. However, in contrast to the NarG clade, each possesses an N-terminal *tat* leader, indicating that each is exported across the cytoplasmic membrane. Catalytic subunits of a number of structurally related derivatives of the CISM architecture are represented in this group. These include catalytic subunits of the *R. sulfidophilum* dimethylamine dehydrogenase (DdhA) [67,68], the *T. selenatis* selenate reductase (SerA), the *I. dechloratans* chlorate reductase (ChlA), and the *Azoarcus* sp. *EB1* ethylbenzene dehydrogenase (EbdA) [39]. The organisms encoding these relatively well-characterized enzymes are all members of the Proteobacteria (a total of 4 Proteobacterial sequences were identified). Sequence alignments and structural data in the case of the EbdA enzyme indicate that these subunits are similar to those of the NarG family in a number of respects: (a) the FS0 cluster is coordinated by one His residue and three Cys residues; (b) the protein–Mo ligand is provided by the carboxylate sidechain of a conserved Asp residue.

A total of 13 sequences were assigned to the DdhA/ClrA/EbdA group. Each of these has 3 residues between C_A and C_B positions of the FS0-coordinating Cys group, with 6 sequences having the His residue at the C_A position. Six sequences share the Asp protein–Mo ligand of NarG, with the ligand in the remaining 7 being unidentifiable or missing from analyses of the sequences. Four sequences share both the C_A His residue and the Asp protein–Mo ligand. Five of the sequences are from *Desulfitobacterium hafniense* (a member of the Firmicutes) with none of these having a readily identifiable protein–Mo ligand. The Firmicutes are additionally represented by a sequence from *Moorella thermoacetica*, which has a Cys at the C_A position and an Asp as protein–Mo ligand. The Archaea are represented by four species: *Haloarcula marismortui* (with a His at the C_A position and an Asp as predicted protein–Mo ligand), *Archaeoglobus fulgidus* (with an Asp as predicted protein–Mo ligand), *Halobacterium salinarium*, and *H. marismortui* (with the final two having a Cys at the C_A position and an unidentified or absent protein–Mo ligand). Overall, the DdhA/SerA/EbdA group has a phylogenic diversity similar to that observed for the NarG clade described above.

Given the use of a 70% identity filter to reduce the size of the sequence dataset studied herein, the appearance of 5 sequences in the DdhA/SerA/EbdA group from *D. hafniense* is remarkable. This organism is a strictly anaerobic member of the Firmicutes that displays a remarkable metabolic diversity, even by the standards of other facultative and obligate anaerobes [158]. This organism appears to be amenable to gene duplication within its genome via the presence of catabolic transposons. As will be seen below in the descriptions of other clades of subunit sequences related to the archetypal CISM catalytic and FCP subunits, the observation of multiple entries for *D. hafniense* is a recurring theme. This organism was originally isolated from sites contaminated with halogenated organic compounds [159–161], suggesting that it would be a good model system for studies of bioremediation [162].

4.1.3. DmsA

These subunits correspond to bona fide catalytic subunits of the CISM archetype DMSO reductase (DmsABC) [78] and form a distinct clade of 24 sequences (Fig. 14). They are distinguished from the soluble periplasmic DMSO and TMAO reductases by the presence of an N-terminal Cys group predicted to coordinate

an FS0 [4Fe-4S] cluster. All members retain a *tat* leader, indicating that they are translocated across the cytoplasmic membrane. As described in Section 2.1, the presence of FS0 has been inferred by comparison with the sequence and structures of the *E. coli* formate dehydrogenase N (FdnGHI) [5,8] and nitrate reductase A (NarGHI) [7,9]. Additional evidence for its occurrence comes from the observation of a [3Fe-4S] cluster by EPR spectroscopy in site-directed mutants [40,41], and the observation of spin–spin interactions between the paramagnetic Mo(V) of the Mo-*bis*PGD cofactor and an adjacent center that can be inferred to be FS0 [62]. The Cys group has a characteristic spacing of three residues between the C_A and C_B positions. The protein–Mo ligand is typically provided by a conserved Ser residue [63]. All the sequences assigned to the DmsA clade have the above characteristics except for one from *D. hafniense* (of a total of 6), which has an undefined or absent protein–Mo ligand.

As is the case for the DdhA/SerA/EbdA group, a large proportion of the identified DmsA sequences are from *D. hafniense* (a member of the Firmicutes)—6 of the 24 identified. An additional Firmicutes sequence is from *M. thermoacetica*. Two sequences are from the *Symbiobacterium thermophilum*, a member of the Actinobacteria which has only been observed in a co-culture with a thermophilic *Bacillus* species [163,164]. The remaining 12 sequences are from the Proteobacteria.

4.1.4. DorA/TorA/BisC

These enzymes are the respiratory periplasmic DMSO and TMAO reductases (DorA and TorA), and the cytoplasmic biotin-*d*-sulfoxide reductase (BisC). The periplasmic enzymes are characterized by the presence of a *tat* leader and a conserved Ser residue as protein–Mo ligand, whereas the cytoplasmic BisC is very similar, except that it lacks the *tat* leader. The DorA/TorA/BisC clade also lacks FS0 and its associated Cys group. Examples of the periplasmic members of this clade have been crystallized and subjected to extensive analyses by X-ray crystallography [14,15,117]. These enzymes have a core structure comprising the 4 domains that form the structural core of the CISM catalytic subunits (Section 2.1).

Thirty-three sequences were identified as being members of the DorA/TorA/BisC clade using the 70% identity cutoff filter. Of these, 26 are targeted for export by a *tat* leader. These sequences are exclusively from members of the Proteobacteria. When the 7 members of the BisC subset are considered, 6 are members of the Proteobacteria, and one is a member of the Actinobacteria (*Rhodococcus RHA1*). Based on this relatively narrow taxonomic distribution, it is tempting to suggest that the entire group of DorA/TorA/BisC enzymes arose at a late stage on the microbial evolutionary timescale.

4.1.5. PgtL

This clade includes the pyrogallol–phloroglucinol transhydroxylase (PgtLS), which is notable for its lack of an FS0 cluster in its catalytic subunit (PgtL) and the lack of an FS4 cluster in its FCP-related TCP subunit (PgtS) [132] (Section 3.1). The protein–Mo ligand is provided by a conserved Ser residue. We identified 5 sequences belonging to the PgtL clade, one from *P. acidigallici* (a member of the Proteobacteria), and 4 from *D. hafniense* (a member of the Fermicutes). None of these sequences bear a *tat* leader.

4.1.6. FdhF

16 Sequences were assigned to the FdhF clade in our analysis. These subunits contain an FS0 [4Fe-4S] cluster that is coordinated by an N-terminal Cys group with two residues separating C_A and C_B, and the protein–Mo ligand is provided by a Sec residue. These features have been confirmed by X-ray crystallography of the FdhF subunit from *E. coli* [34]. As indicated in Section 3, FdhF forms part of the formate hydrogen lyase complex (hydrogenase-3) in combination with a number of additional subunits encoded by the *hyc* operon. Intriguingly, of the 16 sequences assigned to this clade, 5 are fusions of an FS0 and Mo-*bis*PGD binding superdomain to the C-terminus of a HoxU hydrogenase domain predicted to coordinate one [2Fe-2S] cluster and two [4Fe-4S] clusters. This arrangement is reminiscent of the structure of the Nqo3/NuoG subunit of *T. thermophilus* complex I [124,144], except that the latter subunit lacks a Mo-*bis*PGD cofactor.

The link between the CISM catalytic subunit architecture and that of the Nqo3/NuoG subunit of complex I is well established [165,166] (Section 3.3). In a number of bacterial species, including *R. eutropha* (a Proteobacterium), a formate dehydrogenase exists that comprises the following subunits: (i) FdsA, which comprises HoxU and CISM catalytic domains; (ii) FdsB, which contains an additional [4Fe-4S] cluster as well as a flavin (FMN) and a NAD^+-binding site; and (iii) FdsC, which contains an additional [2Fe-2S] cluster [167]. These subunits share considerable structural and functional similarity with the Nqo3/NuoG, Nqo1/NuoF, and Nqo2/NuoE subunits of complex I [124,167]. Thus, it is highly probable that the NAD^+-oxidizing arm of complex I arose from an FdsABC-type formate dehydrogenase, with the Mo-*bis*PGD cofactor being lost during its evolution.

In order to gain further insights into the evolutionary link between the CISM catalytic subunit architecture and the Nqo3/NuoG subunit of complex I, we mined the UNIPROT sequence database using a number of FdsA subunit sequences as bait. These bait sequences were chosen specifically because they contain both the FdhF (CISM catalytic) and HoxU (hydrogenase) domains. We applied an 85% identity cutoff filter to the resultant sequence hits to generate a sequence set of 121 unique sequences. Of these, 81 share the HoxU domain with a FS0 and Mo-*bis*PGD containing domain similar to FdhF that is predicted to have a Sec as the protein–Mo ligand (they are thus homologs of *R. eutropha* FdsA). These originate from the Archaea (4 Crenarchaeota and 3 Euryarchaeota sequences), the Firmicutes (20 sequences), the Actinobacteria (2 sequences), the Deinococcus-Thermus (1 sequence), with the remaining 51 being from the Proteobacteria. The broad phylogenic diversity of these sequences suggests that the FdsA formate dehydrogenase architecture arose early in evolution. An additional group of 19 sequences possess the HoxU and CISM catalytic domains, but lack the Sec protein Mo-ligand, and these sequences are identified as members of the Nqo3/NuoG family in BLASTP

searches. Of this group, 6 are from the Actinobacteria, as well as one from the each of the Nitrospirae, Chlamydiaea, Acidobacteria, and Firmicutes. The remaining 9 are from the Proteobacteria. None of these sequences are Archaeal in origin, which is consistent with the observation that the complete complex I architecture is absent from this domain of life [168]. Finally, a group of 12 sequences very closely related to the FdhF subunit was identified (containing only the CISM catalytic subunit domain). Of this final group, two are from the Firmicutes and 10 are from the Proteobacteria. Thus, sequence analyses of the FdhF clade of CISM catalytic subunits can provide important insights into the evolution of components of complex I that complement the structural approaches described in Section 3. It should be noted, however that the FdhF clade of Fig. 14 is not adjacent to that of the Nqo3/NuoG clade. This probably reflects a divergence of sequence conservation between the two that corresponds to residues surrounding the Mo-*bis*PGD binding pocket of the FdhF clade and the vestigial binding pocket of Nqo3/NuoG clade.

4.1.7. NapA

Eight sequences were identified as members of the NapA clade. These include examples of soluble periplasmic (*tat*-translocated) and cytoplasmic nitrate reductases [169,170]. Typically, the periplasmic members of this clade form heterodimers with a NapB diheme cytochrome *c* (Table 2), whereas the cytoplasmically localized examples are monomeric. This clade contains members of the Proteobacteria (3 sequences), Cyanobacteria (2 sequences), Actinobacteria, (1 sequence), and the Firmicutes (1 sequence). Members of the NapA clade are characterized by the presence of an FS0 [4Fe-4S] cluster and a conserved Cys residue as protein–Mo ligand.

4.1.8. FdnG and W-FdhA

Inspection of Fig. 14 reveals that a single clade has members that are functionally assigned as either catalytic subunits of CISM archetype FdnGHI enzymes or catalytic subunits of the dimeric tungsten-containing formate dehydrogenase of *D. gigas* [120]. In both cases, the subunits are predicted to contain an FS0 cluster, a conserved Sec protein–Mo ligand, and are predicted to be targeted to the periplasmic compartment by the presence of an N-terminal *tat* leader. The FdnG subgroup contains 12 sequences, from the Proteobacteria (8 sequences), the Chloroflexi (1 sequence), the Aquificae (1 sequence), and one from an undefined environmental sample. The W-FdhA subgroup comprises 9 sequences, seven from the Proteobacteria, one from the Actinobacteria and one from the Firmicutes. In addition to the differences in subunit composition between the FdnG and W-FdhA groups, sequences from the latter group appear to be largely from thermophilic species.

4.1.9. NuoG/Nqo3/AoxB

This clade contains the few Nqo3/NuoG (7 sequences) hits returned using the archetypal CISM catalytic subunits as bait. As discussed above, the Nqo3/NuoG sequences appear to have arisen from a gene fusion event involving a HoxU-type subunit. None of these sequences has an identifiable protein–Mo ligand, and in the complex I from *T. thermophilus*, an empty (water filled) cavity exists in a location equivalent to the Mo-*bis*PGD of the archetypal CISM catalytic subunit [124]. An additional sequence assigned to this group is the catalytic subunit of *Cenibacterium arsenoxidans* arsenite oxidase (AoxB) [119]. This subunit lacks a protein–Mo ligand and has a [3Fe-4S] cluster as its FS0 cluster, which is coordinated by an incomplete Cys group lacking the fourth Cys residue (C_D). Additionally, it forms a functional dimer with a subunit (AoxA) containing a Rieske-type [2Fe-2S] cluster that directs the complex to the periplasmic compartment via a *tat* leader at its N-terminus. Members of the Nqo3/NuoG clade have a fairly broad taxonomic distribution, but are lacking from the Archaea (4 Proteobacteria, 1 Deinococcus-Thermus, and 2 Actinobacteria).

4.1.10. Undefined-1

Clear functional assignments could not be made for two clades of Fig. 14 based on BLASTP searches. The first such clade, labeled as Undefined-1 in Fig. 14, has 13 members that are broadly similar to the catalytic subunits of the PhsABC, PsrABC, and DmsABC CISM archetypes. Nine sequences have a *tat* leader, whereas 4 do not. All are predicted to coordinate a [4Fe-4S] FS0 cluster and with one exception, they each have a Cys residue as protein–Mo ligand (the exception being undefined). Six sequences are from the Proteobacteria, 2 from the Chlorobi, 2 from the Firmicutes, 2 from the Actinobacteria, and 1 from the Chlorflexi.

4.1.11. TtrA

23 Sequences were assigned to the tetrathionate reductase catalytic subunit clade: 14 were from the Proteobacteria, 4 from the Firmicutes, 2 from the Chlorflexi, and 3 from the Archaea. Based on this taxonomic range, it is likely that the TtrABC architecture arose early in evolution. Each sequence bears an N-terminal *tat* leader, is predicted to coordinate an FS0 [4Fe-4S] cluster, and to have a Cys residue as its protein–Mo ligand. The active site of TtrABC catalyzes the reduction of tetrathionate to thiosulfate [114,171], which is itself reduced to sulfide and sulfite by thiosulfate reductase (PhsABC) (see below).

4.1.12. PhsA (without a tat leader)

9 Sequences appear in a clade whose members are catalytic subunits of the thiosulfate reductase CISM catalytic subunits. Each sequence lacks a *tat* leader, is predicted to coordinate an FS0 [4Fe-4S] cluster, and to have a Cys protein–Mo ligand. All sequences are from the Proteobacteria, implying their appearance at a fairly late stage of evolution. A thiosulfate reductase with a putative cytoplasmic catalytic subunit would require the presence of a thiosulfate transport system, such as that encoded by the *cysPTWA*-type operons of *E. coli* and *S. typhimurium* [172,173]. BLASTP searches using the sequence of the periplasmic thiosulfate binding protein (CysP) as bait reveal that thiosulfate uptake systems are widespread amongst the Proteobacteria.

4.1.13. PhsA/PsrA (with a tat leader)

23 Sequences appear in a clade with individual members being assigned as catalytic subunits of polysulfide or thiosulfate

reductases. These PsrABC and PhsABC enzymes are CISM archetypes. Each member of this clade has an N-terminal *tat* leader, is predicted to coordinate a FS0 [4Fe-4S] cluster, and to have a Cys residue as protein–Mo ligand. 12 Members of this clade are Proteobacteria, 3 are Firmicutes, 3 are Chlorobi, one is a Deinococcus-Thermus, and 4 are from the Archaea. This relatively broad taxonomic range implies the emergence of this overall class at a relatively early stage of evolution.

4.1.14. Undefined-2

The final clade of CISM catalytic subunits comprises 28 sequences that are variously identified as homologs of the *E. coli* DmsA paralog YnfE, tungsten containing formate dehydrogenase (W-FdhA), PhsA, DdhA (dimethylamine dehydrogenase), and SerA (selenate reductase). With the exception of two YnfE homologs from *D. hafniense*, none of these sequences bears a *tat* leader, and all appear to have a Cys residue as protein–Mo ligand. Ten of the sequences are from *D. hafniense*. Interestingly, 5 sequences, 4 from *D. hafniense* and 1 from *Desulfuromonas acetoxidans* (a Proteobacterium) have an incomplete Cys group lacking a C_A. 14 Sequences are from the Proteobacteria, 11 from the Firmicutes (including 10 from *D. hafniense*), and 2 are from the Planctomycetes (*Candidatus Kuenenia stuttgartiensis*).

A final novel feature of this clade is that 3 sequences (from *C. Kuenenia stuttgartiensis* (Planctomycetes), *Magnetococcus* sp. MC-1 (Proteobacterium), and *Geobacter* sp. FRC-32 (Proteobacterium)) have a C-terminal NifS domain (Cys desulfurase). Also, two Proteobacteria sequences (from *R. eutropha* and *Burkholderia cepacia*) have C-terminal fusions to a domain resembling the FAD-binding domain and NADH-binding site of nitric oxide dioxygenase (HMP) [174].

4.2. The FCP subunits

Fig. 15 illustrates the phylogenic relationships between the FCP subunit sequences obtained as described above. Because of the more general function of the FCP subunits in ETR delineation between active sites in comparison with the varied substrate specificities of the proteins related to the CISM catalytic subunits, fewer distinct clades are observed for sequences identified as being FCP subunits (9 versus 15). Members of the individual clades are summarized as follows:

4.2.1. NarH

As described in Section 2.2, a typical NarH subunit comprises three blocks of core sequence that provide coordination for the FS1–FS4 clusters of the ETR, along with three additional surface-localized subdomains. A *tat* leader is absent, and the third Cys group (Group III) has a Trp residue at the C_B position. The NarH clade of Fig. 15 contains 28 sequences that share these characteristics. Not surprisingly, the taxonomic diversity of NarH sequences reflects that observed for NarG, with 2 Archaea, 12 Actinobacteria, 5 Firmicutes, one Deinococcus-Thermus, and 8 Proteobacteria being represented. As is the case with the NarG clade of Fig. 14, this taxonomic range supports the assertion that the NarGHI architecture likely arose at an early stage of evolution. None of the NarH sequences bears an N-terminal *tat* leader.

4.2.2. DdhB/SerB/EbdB

As is the case for the catalytic subunit clade comprising DdhA, SerB, and EbdA sequences, this clade branches off the NarH clade and is closely related to it. Five sequences were identified as members of the DdhB/SerB/EbdB, one from the

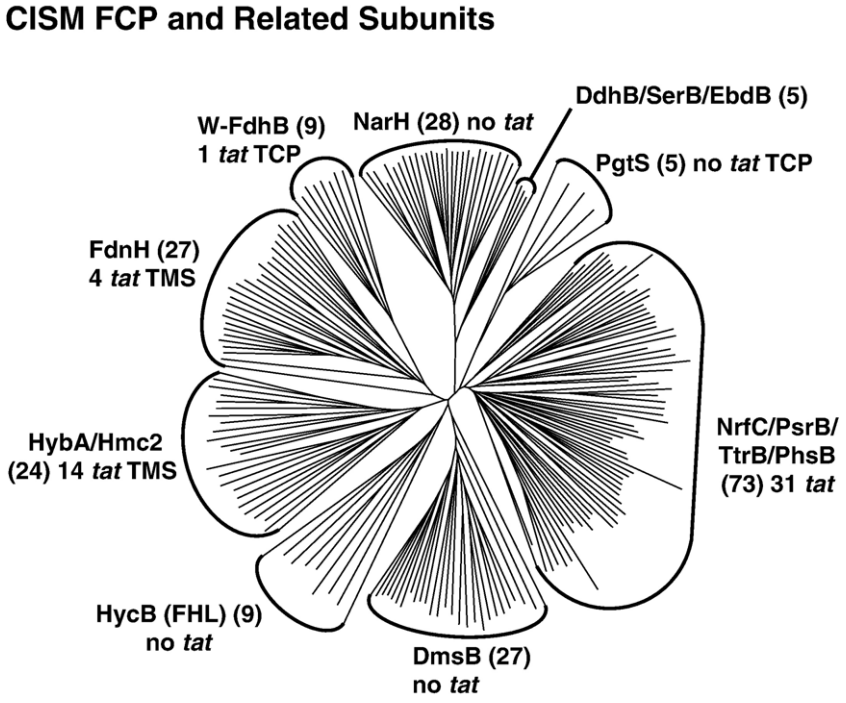

Fig. 15. Phylogenic analysis of the FCP and related subunits. Sequences were mined and analyzed as described in the text.

Archaea and the remaining 4 from the Proteobacteria. Each of these is predicted to coordinate three [4Fe-4S] clusters (FS1, FS2, and FS3) and one [3Fe-4S] cluster (FS4). The predicted presence of a [3Fe-4S] cluster is explained by the presence of a His, Tyr, or Ala at the C_B position of Cys group III. These predictions are supported by EPR characterization of the DdhABC enzyme from *R. sulfidophilum* [68] and the determination of the structure of EbdB subunit of ethylbenzene dehydrogenase from *A. aromaticum* [39]. None of the identified DdhB/SerB/EbdB sequences bears an N-terminal *tat* leader.

4.2.3. PgtS

This clade has members that are identified on the basis of BLASTP searches as being TCP subunits of pyrogallol–phloroglucinol transhydroxylase-type enzymes. As is the case for the PgtL clade of Fig. 14, the identified PgtS sequences of Fig. 15 arise from two organisms: one from *P. acidigallici* (a Proteobacterium) and 4 from *D. hafniense* (a Firmicute). Members of this clade have an incomplete Cys group III which lacks $C_{A–C}$, as well as lacking C_D from Cys group II, resulting in an absence of the FS4 cluster [132]. As indicated above, the PgtLS enzyme also lacks an FS0 cluster in its catalytic subunit (PgtL). As mentioned in Section 3.1, it is possible that the PgtLS dimer is anchored to the cytoplasmic membrane by a C-terminal domain with a fold similar to that of the cell-adhesion protein fibronectin III [133].

4.2.4. NrfC/PsrB/TtrB/PhsB

This clade contains a large number of sequences (73) that are identified as being members of the NrfC, PsrB, TtrB, and PhsB families on the basis of BLASTP searches or their annotations in the SWISSPROT database. Each protein sequence is predicted to coordinate 4 [4Fe-4S] clusters. Of these, those assigned as being NrfC sequences are predicted to be directed to the *tat* translocon by an N-terminal leader sequence. Given the large number of sequences assigned to this clade (based on the 70% identity filter), it is not surprising that a wide phylogenic range is represented: 44 sequences were from the Proteobacteria, 12 were from the Firmicutes, 1 from the Aquificae, 6 from the Archaea, 5 from the Chlorobi, 1 from the Chloroflexi, 2 from the Actinobacteria, 1 from the Deinococcus-Thermus, and 1 from the Acidobacteria. The wide taxonomic range observed in this clade is a reflection of the utility of the FCP architecture in defining effective ETR systems in a range of enzyme systems. A total of 31 sequences within this clade have an N-terminal *tat* leader.

4.2.5. DmsB

27 Sequences were identified as being members of the DmsB clade. Each of these is predicted to coordinate four [Fe-S] clusters, FS1–FS4. These clusters are predicted to be [4Fe-4S] clusters in every sequence except one from the Firmicute *D. hafniense*, in which the C_B position of Cys group III is occupied by an Ala, resulting in a predicted cluster composition of three [4Fe-4S] clusters and one 4 [3Fe-4S] cluster. Of the 27 sequences, 13 are from the Firmicutes, 4 from the Actinobacteria, and 10 from the Proteobacteria. This distribution is similar to that observed for the DmsA catalytic subunit clade. 11 of the 13 Firmicutes sequences are from *D. hafniense*, with the remaining two from *M. thermoacetica*. The three Actinobacteria sequences are from *S. thermophilum*. None of the sequences bears a *tat* leader.

4.2.6. HycB

9 Sequences were assigned to this clade, 2 of which are from the Archaea, with the remainder being from the Proteobacteria. HycB functions as part of the *E. coli* FHL complex as an electron transfer module, linking the FdhF subunit (formate dehydrogenase) to the hydrogenase-type subunits of the complex (hydrogenase-3). One of the sequences in this clade encodes an FCP subunit of a carbon monoxide dehydrogenase from *Rhodospirillum rubrum* [175].

4.2.7. HybA/Hmc2

24 Sequences were assigned to this clade. Functionally, these sequences appear to comprise FCP subunits of the electron donor system in the hydrogenase system corresponding to hydrogenase-2 of *E. coli* or the periplasmic hydrogenase of *D. vulgaris*. In addition to coordinating four [4Fe-4S] clusters, these sequences each have a TM segment located towards their C-termini (predicted using the EMBOSS tmap program [176]). Fourteen of the sequences bear an N-terminal *tat* leader. 4 Sequences within this clade are from the Firmicutes, 1 from the Chloroflexi, 1 from the Actinobacteria, 5 from the Acidobacteria, and 13 from the Proteobacteria.

4.2.8. FdnH

27 Sequences appear in a well-defined clade with members being assigned as FdnH sequences on the basis of BLASTP searches. Like the HybA/Hmc2 clade, these sequences are predicted to have a C-terminal TM segment. A *tat* leader is lacking in all except 4 of the sequences. Sequences were identified from 1 Archaea, 4 Firmicutes, 4 Actinobacteria, 1 Aquificae, 1 Acidobacteria, and 13 Proteobacteria. An additional sequence was from an undefined environmental sample. This taxonomic distribution is slightly wider than that observed for the FdnG catalytic subunit clade (Fig. 14). The FdnH clade includes an Archaeal member (from *P. aerophilum*) that is lacking from the FdnG clade, as well as members of the Firmicutes. Four of the 27 members of this clade are predicted to have a *tat* leader.

4.2.9. W-FdhB

Nine sequences were identified as being electron transfer subunits of tungsten-containing formate dehydrogenases (W-FdhAB) such as that found in *D. gigas*. By virtue of their incomplete Cys group III, which lacks C_A–C_C, and a Cys group II which lacks C_D, these subunits are predicted to coordinate three [4Fe-4S] clusters and are thus deemed to be TCP subunits that lack FS4. All 9 sequences are from the Proteobacteria, with 7 from *Desulfovibrio* strains, and two from *Syntrophobacter fumaroxidans*. As is the case for the FdnG and FdnH sequences, a strict correlation does not exist between the taxonomic distributions of the W-FdhA and W-FdhB sequences. Also, it is notable that the two clades of TCP sequences (PgtS and W-FdhB) are not adjacent in Fig. 15, suggesting that the TCP architecture may have arisen at least twice during the evolution of the bacterial species for which sequence information is currently available

(convergent evolution). Eight of the nine sequences lack a *tat* leader.

5. Soluble periplasmic molybdoenzymes and bacterial cell envelope morphology

As indicated in Section 3, soluble periplasmic molybdoenzymes that share the CISM catalytic subunit architecture can exist as monomers, or they can pair with an additional subunit to form soluble dimers. In both cases, the Mo-*bis*PGD containing subunit is targeted to the periplasmic compartment by the *tat* translocon. The large amount of sequence data on the CISM enzymes and relatives presents an excellent opportunity to investigate if there is a correlation between the existence of soluble exported Mo-*bis*PGD-containing enzymes and the presence of an outer membrane. To address this, we used a subset of subunits of the enzymes as described in Table 2, some of which are soluble and others of which are membrane associated in multisubunit complexes. The results of our analyses are summarized in Table 4. Because a significant number of species possessing outer membranes are reported as lacking one on the basis of the Gram stain, we are reluctant to use the "Gram positive"/"Gram negative" terminology herein. Our assignment of the presence of an outer membrane is based on reported morphological studies of representative members of individual bacterial phyla.

In order to test our hypothesis, we chose a number of soluble Mo-*bis*PGD containing subunit families whose members are either located in the periplasm (NapA and TorA/DorA), or the cytoplasm (NarB and BisC). As controls, we used molybdoenzyme catalytic subunits that are parts of membrane-bound multisubunit complexes, including the CISM archetypal catalytic subunits NarG (of NarGHI), DmsA (of DmsABC), PsrA/PhsA (of polysulfide reductase, PsrABC, and thiosulfate reductase, PhsABC). Of these control subunits, NarG is clearly located on the inside of the cytoplasmic membrane and PsrA and PhsA are located on the periplasmic side of the cytoplasmic membrane. As described in Section 2.3.3., some controversy continues to surround the actual location of the DmsA subunit of the DmsABC complex, although it is likely to be attached to the periplasmic side of the membrane. Sequences were mined from the UNIPROT sequence database (http://www.ebi.ac.uk/UNIPROT/index.html) and hits were filtered by pairwise comparison with a 95% identity cutoff filter. The presence of a *tat* leader was deemed to result in a periplasmic functional location for all the proteins analyzed.

The NapA and NarB families were selected because members of the former are *tat*-targeted periplasmic free floaters (as dimers with NapB subunits), whereas members of the latter are soluble and cytoplasmically localized. These enzymes participate in dissimilatory and assimilatory nitrate reduction, respectively, and are predicted on the basis of sequence alignments to have essentially identical overall structures. In the case of NapA, our dataset comprises 85 sequences, of which 83 are from bacteria possessing an outer membrane, indicating that there is a strong bias against finding *tat*-targeted NapA sequences in bacterial species lacking an outer membrane (a ratio of 41.5:1). The two organisms identified that appear to have a soluble exported NapA-type nitrate reductase are *S. thermophilum* [163,164] and *D. hafniense* [158]. It is not clear why these species appear to export NapA-type subunits, but the former has only been observed in a co-culture with a thermophilic *Bacillus* strain, and the latter is notable for the large number of redox enzymes

Table 4
Correlation between soluble periplasmic molybdoenzymes and the presence of an outer membrane

Catalytic subunit family	Soluble[a]	Part of membrane-bound complex?	*Tat* leader?[b]	Total sequences (UNIPROT database)[c]	Species with outer membrane[d]	Species without outer membrane[d]	Ratio[e]
NapA (periplasmic nitrate reductase)	Yes	No	Yes	85	83	2	41.5:1
NarB (cytoplasmic nitrate reductase)	Yes	No	No	36	28	8	3.5:1
TorA/DorA (periplasmic DMSO/TMAO reductase)	Yes	No	Yes	62	62	0	∞:1
BisC (cytoplasmic biotin-*d*-sulfoxide reductase)	Yes	No	No	36	27	9	3:1
NarG (membrane-bound nitrate reductase A catalytic subunit)	No	Yes	No	97	58	39	1.5:1
DmsA (membrane-bound DMSO reductase)	No	Yes	Yes	41	33	8	4.1:1
PsrA/PhsA (membrane-bound polysulfide reductase—PsrA; thiosulfate reductase—PhsA)	No	Yes	Yes	40	33	7	4.7:1
Complete bacterial genomes[f]	–	–	–	–	300	177	1.7:1

[a] Soluble enzymes are defined as those which are not part of a membrane-bound complex.
[b] The *Tat* leader sequence was identified by inspection of ClustalX sequence alignments.
[c] BlastP searches were used to identify sequences belonging to the individual catalytic subunit families, and these were retrieved from the UNIPROT sequence database. To avoid duplication, the sequence sets were filtered by pairwise comparison to remove sequences with greater than 95% sequence identity to any other sequence in each set. Multiple sequence entries from single species were treated as single entries for the purposes of this analysis. Sequence mining was completed in March 2007.
[d] To avoid confusion, we do not use the "Gram positive"/"Gram negative" nomenclature. Species with outer membranes were identified on the basis of morphological studies or the presence of readily identifiable outer membrane proteins encoded by their genomes. The following domains/phyla were deemed to lack outer membranes: Archaea, Fermicutes, Actinobacteria, and Planctomycetes. The following bacterial phyla were deemed to have outer membranes: Deinococcus-Thermus, Chloroflexi, Cyanobacteria, Spirochetes, Bacteroidetes, Chlamydiae, Aquificae, Acidobacteria, Fusobacteria, Thermotogae; and Nitrospina.
[e] Ratio of number of sequences from strains having an outer membrane to the number of sequences from strains lacking an outer membrane.
[f] Summary data obtained from www.genomesonline.org, and the presence of an outer membrane was determined as described in Footnote d.

encoded by its genome. The cytoplasmic NarB dataset comprises 36 sequences, of which 28 are from species having an outer membrane and 8 were from species lacking one, a ratio of 3.5:1. Overall, analyses of the NapA and NarB families of catalytic subunits support the hypothesis that exported soluble molybdoenzymes are largely confined to bacteria possessing an outer membrane.

The TorA/DorA and BisC families mimic the differences observed between the NapA and NarB families. Both are soluble with the former family being *tat*-targeted to the periplasmic compartment and the latter being cytoplasmic. They catalyze respiratory *N*- or *S*-oxide reduction and biotin-*d*-sulfoxide reduction, respectively. In the case of the TorA/DorA TMAO/DMSO reductases, 62 sequences were identified, and all of these are from bacteria possessing an outer membrane. The closely related cytoplasmic BisC family comprises 36 sequences, of which 27 are from organisms possessing an outer membrane and 9 are from organisms lacking one, a ratio of 3:1.

In the case of the control subunits from CISM archetypes, we anticipated that there would be a much greater representation of sequences from organisms lacking an outer membrane. 97 NarG sequences were obtained from UNIPROT, 58 of which were from species possessing an outer membrane and 39 from species lacking one, a ratio of 1.5:1. 41 DmsA sequences were obtained, 33 of which were from species possessing an outer membrane, with 8 being from species lacking one, a ratio of 4.1:1. Finally, we analyzed the PsrA/PhsA polysulfide/thiosulfate reductase family of catalytic subunits. In this case 40 sequences were obtained, of which 33 are from species possessing an outer membrane and 7 from species lacking one, a ratio of 4.7:1.

To estimate the expected ratio between protein sequences from strains having an outer membrane to those lacking one, we analyzed the taxonomic information contained in the Genomes Online Database (www.genomesonline.org). As of March 2007, the sequences of 477 complete bacterial genomes have been reported. Of these, 300 are of species with outer membranes, and 177 are of species lacking an outer membrane, a ratio of 1.7:1 (see legend to Table 3 for details). Of the subunits chosen for analysis, only the distribution of the NarG subunits between the two primary bacterial morphologies is close to that predicted (a ratio of 1.5:1). That this is not observed in the other subunit families, such as the BisC or the DmsA/PsrA/PhsC families is likely to be a consequence of a combination of the evolution of the metabolic pathways involved and evolutionary emergence of the outer membrane itself. Overall, our analysis has clearly demonstrated a bias against the presence of soluble periplasmic molybdoenzymes in species lacking an outer membrane.

6. Evolutionary relationships between the catalytic/FCP subunits and other redox systems

6.1. Relationships between the catalytic subunit architecture and other systems

The complexity of the CISM catalytic architecture, with its complex non-contiguous domain structure, appears to confound facile attempts to understand how it arose, presumably at a very early stage of evolution in the Archaeal domain of life [115]. It is likely the result of a complex series of primordial gene fusion and duplication events. However, as outlined in Section 4.1, the basic architecture can be modified to accommodate a very wide range of water soluble substrate interconversions, ranging from the relatively mundane reduction of compounds such as nitrate to reactions that do not involve a net redox reaction such as the interconversion of pyrogallol and phloroglucinol by *P. acidigallici* PgtAB [130,131]. Comparison of the structure of the CISM catalytic subunit with that of the Nqo3/NuoG subunit of Complex I reveals a likely evolutionary relationship between these two families of proteins. The more conventional bioinformatic analyses presented in Section 4.1 reveal that some members of the FdhF clade of catalytic subunits are fused to the C-terminus of a HoxU-type hydrogenase domain, suggesting that the Nqo3/NuoG architecture is derived from such a fusion, after which the Mo-*bis*PGD binding functionality was lost. Thus, although the origin of the catalytic subunit architecture is unclear, its multifunctional nature has subsequently been exploited by a wide variety of species.

Consideration of the phylogenic distribution of catalytic subunit families containing an FS0 cluster (NarG, for example) with those lacking such a cluster (TorA, for example) reveals that the former group has a wider distribution than the latter (Section 4.1). Thus, it is likely that the four-domain TorA/DorA architecture arose as a simplification of the more complex FS0-containing architecture.

6.2. Evolution of the FCP architecture

As indicated above, the archetypal FCP subunit contains four tetranuclear [Fe-S] clusters coordinated by Cys groups similar to those observed in the bacterial 2[4Fe-4S] ferredoxins (Section 2.2 and Fig. 6). The evolution of the bacterial 2[4Fe-4S] ferredoxins has been reviewed elsewhere [177,178] and will not be covered in detail herein. Briefly, it is believed that the original 2[4Fe-4S] cluster binding domain arose from the interaction of two more-or-less identical peptides, each containing 4 Cys residues. Three Cys residues from one peptide and one from the second coordinates one [4Fe-4S] cluster, whilst the reverse applies to the other cluster [179]. Eventually, the genes encoding the two Cys groups became fused, and the resultant protein loosely resembles the subdomain of the FCP architecture that coordinates FS3 and FS4 (Cys groups II and III, Fig. 6). Subsequent evolution resulted in the loss of, for example, one Cys at one of the C_B positions to generate a protein coordinating a [4Fe-4S] cluster and a [3Fe-4S] cluster, as is found in the 7Fe ferredoxin from *T. thermophilus* [180]. Additionally, loss of multiple Cys residues resulted in the complete absence of one of the clusters, as is the case in the [4Fe-4S] cluster ferredoxin from *D. africanus* [181]. It is evident that many of the genetic events leading to the various architectures observed in the bacterial ferredoxins arose at an early juncture in evolution [179,182].

To gain insights into the evolution of the FCP architecture, we extended the SSM approach described in Section 3.1. Instead of searching for structural similarity over the entire C-α structure of the bait used (NarH and FdnH), we allowed

additional searches for subdomains to be performed within the PDB archive. Table 5 summarizes the results of these analyses. Not surprisingly, structural alignments are identified between both bait subunits and the PgtS subunit from *P. acidigallici* and the W-FdhB subunit of *D. gigas*. Both these hits are TCP subunits lacking FS4. In both cases, hits are observed to two bacterial ferredoxins, the 2[4Fe-4S] cluster ferredoxin from *Clostridium acidi-urici* (PDB code 1FDN) [183] and the [4Fe-4S] cluster ferredoxin from *Desulfovibrio africanus* [181]. In the case of NarH, these hits occur with rmsd values of 1.71 Å and 2.27 Å over 49 and 50 residues, respectively. Very similar results were obtained using the FdnH structure as bait (Table 5).

Fig. 16A shows the structural alignments identified by the SSM server using the PYMOL molecular graphics package. The alignments shown are with the FdnH subunit, but essentially similar results are obtained when they are performed with NarH. Fig. 16B shows the structurally aligned sequences separately and illustrates the absence of the cluster corresponding to FS2 in the structure of the *D. africanus* ferredoxin (1FXR). Inspection of the sequence of this protein reveals degradation of the Cys group corresponding to Cys group IV of the FCP architecture. The second ferredoxin structure, from *C. acidi-urici* (1FDN), retains two complete Cys groups that are structurally equivalent to Cys groups II and III of the FCP architecture. Fig. 17 shows a graphical representation of the structural alignments, indicating that the *D. africanus* [4Fe-4S] ferredoxin structurally aligns with protein sequences encompassing Cys groups I and IV, whereas the *C. acidi-urici* 2[4Fe-4S] ferredoxin structurally aligns with Cys groups II and III.

Table 5
SSM search results using the structure of the NarH and FdnH as bait

Hit: subunit and PDB code	System type	rmsd (Å)	N_{align} [a]
To NarH (1Q16)			
PgtS TCP subunit (*P. acidigallici*) pyrogallol–phloroglucinol transhydroxylase (1TI2)	Catalytic-TCP dimer	1.73	176
FdhB TCP subunit (*D. gigas*) W-containing formate dehydrogenase (1H0H)	Catalytic-TCP dimer	1.92	156
FdnH FCP subunit (*E. coli*) (1KQF)	CISM archetype	1.75	157
2[4Fe-4S] ferredoxin (*C. acidi-urici*) (1FDN)	Bacterial 2[4Fe-4S] ferredoxin	1.71	49
[4Fe-4S] ferredoxin (*D. africanus*) (1FXR)	Bacterial [4Fe-4S] ferredoxin	2.27	50
To FdnH (1KQF)			
FdhB TCP subunit (*D. gigas*) W-containing formate dehydrogenase (1H0H)	Catalytic-TCP dimer	1.48	203
PgtS TCP subunit (*P. acidigallici*) pyrogallol–phloroglucinol transhydroxylase (1TI2)	Catalytic-TCP dimer	2.05	166
NarH (*E. coli*) NarGHI (1Q16)	CISM archetype	1.68	155
[4Fe-4S] ferredoxin (*D. africanus*) (1FXR)	Bacterial [4Fe-4S] ferredoxin	2.18	59
2[4Fe-4S] ferredoxin (*C. acidi-urici*) (1FDN)	Bacterial 2[4Fe-4S] ferredoxin	2.50	50

[a] Number of residues aligning with quoted rmsd, out of a total of 509 residues for the NarH subunit and 289 residues for the FdnH subunit.

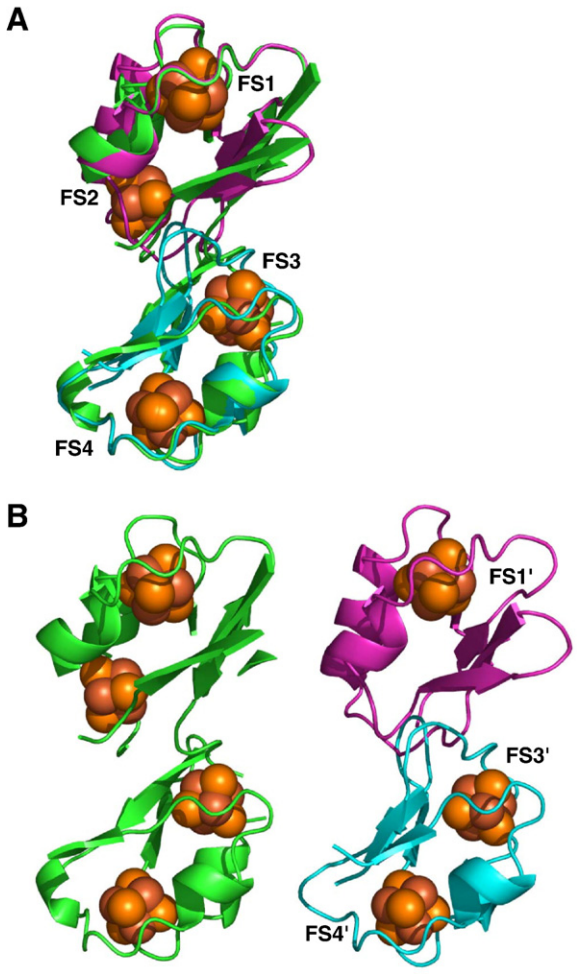

Fig. 16. Structural alignments of FdnH (1KQF, green) with the [4Fe-4S] ferredoxin from *Desulfovibrio africanus* (1FXR, pink) and the 2[4Fe-4S] ferredoxin from *Clostridium acidi-urici* (1FDN, blue). (A) Aligned structures. (B) Offset structures. Note the absence of FS2' in the structure of the *Desulfovibrio africanus* [4Fe-4S] ferredoxin. The alignments with the B chain of 1KQF (FdnH) can be summarized as follows: 1FXR aligns with an rmsd of 2.2 Å over 64 C-α atoms, whereas 1FDN aligns with an rmsd of 2.5 Å over 55 C-α residues. In both panels, non-aligned structure is omitted for clarity.

Similar analyses have been carried out by Kloer et al. [39] on the FCP subunit structure of *Azoarcus* sp. *EB1* ethylbenzene dehydrogenase (EbdABC), clearly supporting the conclusion that the FCP architecture contains structural elements closely related to the bacterial 2[4Fe-4S] ferredoxins.

The lack of a cluster equivalent to FS2 in the *D. africanus* ferredoxin (1FXR) is intriguing. The explanation for this probably lies in the established evolutionary pathway for the loss of the second [4Fe-4S] cluster in the bacterial ferredoxins to generate either an empty pocket, or a pocket containing an alternative cluster such as a [3Fe-4S] cluster [177,178]. To further investigate this, we used the 1FXR structure as bait to search the entire PDB database using the SSM server. This search returned a [3Fe-4S] [4Fe-4S] ferredoxin from *T. thermophilus* (PDB code 1H98) [180], which aligned with the 1FXR structure with an rmsd of 1.9 Å over 53 C-α atoms (Fig. 18). Overall, these structural alignments with the bacterial ferredoxins provide important clues as to the evolutionary pathway leading to the FCP architecture

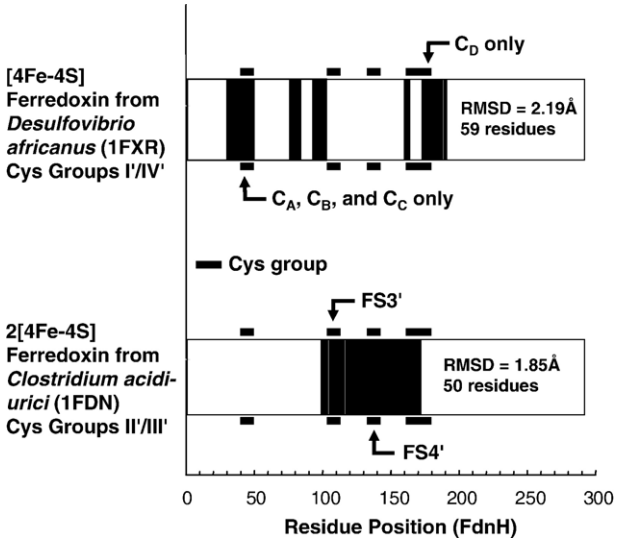

Fig. 17. Graphical representation of the structural alignment between FdnH (1KQF) and the [4Fe-4S] ferredoxin from *Desulfovibrio africanus* (1FXR) and the 2[4Fe-4S] ferredoxin from *Clostridium acidi-urici* (1FDN). The positions of the reported structural overlaps are indicated in black. The position of the Cys groups of FdnH are as indicated by the four bars above and below the alignment.

that complement similar conclusions derived from more conventional sequence-based bioinformatics.

So what is the evolutionary pathway leading to the FCP architecture? Fig. 19A shows a possible route leading from a gene insertion event in which the gene of one 2[4Fe-4S] ferredoxin is inserted between the two Cys groups of another. Such an event would result in a "primordial" FCP. Fig. 19B shows the consequences of such a gene fusion at the protein level, with the clusters of one ferredoxin being equivalent to FS4 and FS3, and

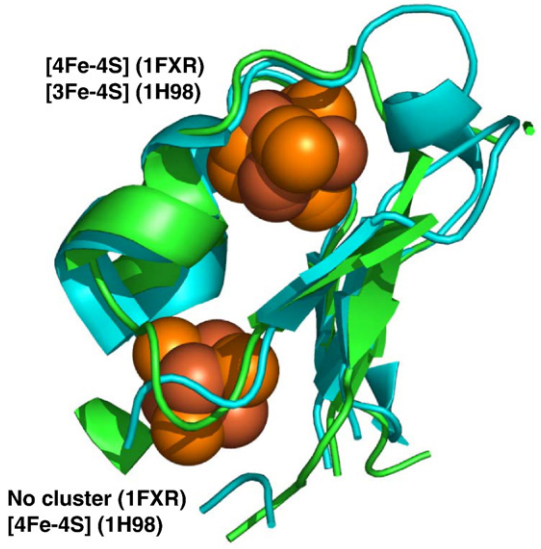

Fig. 18. Structural alignment of the [4Fe-4S] ferredoxin from *Desulfovibrio africanus* (1FXR) with the [3Fe-4S] [4Fe-4S] ferredoxin from *Thermus thermophilus* (1H98). Secondary structures corresponding to aligned C-α atoms are shown. Green: the [4Fe-4S] ferredoxin from *D. africanus*. Blue: the [3Fe-4S][4Fe-4S] ferredoxin from *T. thermophilus*. The two structures align with an rmsd of 1.9Å over 53 C-α atoms.

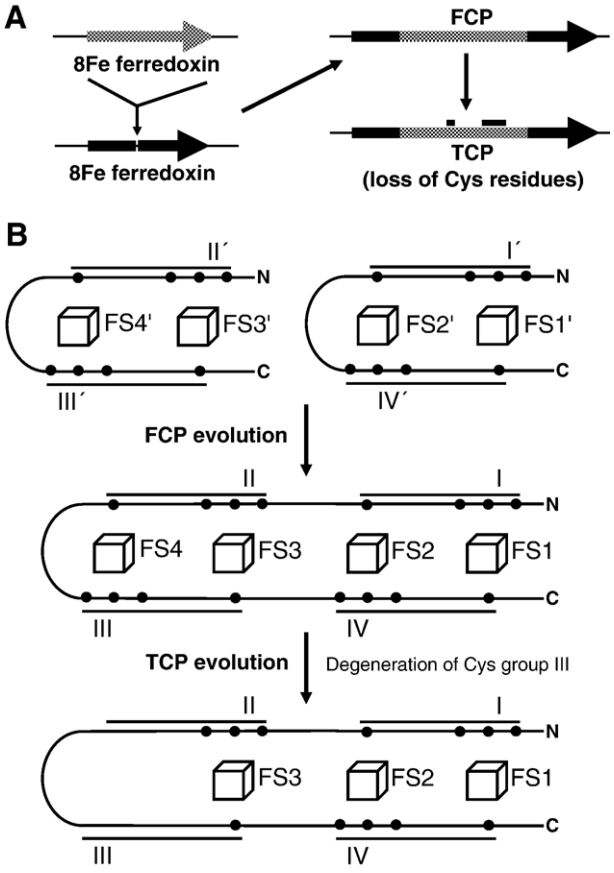

Fig. 19. Evolution of the FCP architecture via the insertion of one 2[4Fe-4S] cluster gene into another. (A) Gene insertion of one 2[4Fe-4S] cluster ferredoxin gene between the two Cys groups of another. (B) Consequences of the gene insertion event for the encoded proteins. Also illustrated is the genetic degradation of Cys group III, resulting the in the TCP architecture of PgtS and W-FdhB.

those of the other being equivalent to FS2 and FS1. Such a gene fusion would account for the observation of a linear sequence of clusters found in all structures of FCP subunits observed to date. The generation of the TCP architecture (PgtS and W-FdhB) would result from a process similar to that which results in the loss of single clusters from the 2[4Fe-4S] ferredoxins: namely the progressive loss of the Cys residues for coordinating FS4 (Cys group III). Thus, in contrast to the situation with the CISM catalytic subunit architecture, it is possible to trace a plausible evolutionary route from the primordial proteins comprising Cys group containing homodimers all the way to the FCP architecture and on to that of the TCP type of subunit.

7. Conclusions and outlook

The explosion of bioinformatic data in recent years has been complemented by an abundance of structural data that enables emerging bioinformatic approaches to be used to gain new insights into the CISM archetypes and closely related proteins. In this review, we have attempted to establish the importance of the CISM subunit architectures, particularly those of the catalytic and FCP subunits, in being responsible to a significant extent for the remarkable metabolic diversity of bacteria ranging

from "primordial" Archaea to the relatively recently emerged Proteobacteria. The recent growth of the PDB database has enabled us to apply the emerging techniques of pairwise SSM analyses and searches of the entire PDB database. We predict that in the near future such searches will become as important as the more typical sequence-based bioinformatics techniques have become in the last decade or so. The recent exponential growth of completed genome sequences has also enabled us to carry out large-scale analyses of the relatedness of CISM catalytic and FCP subunits, as well as close functional relatives. We have been able to exploit this sequence information to demonstrate a clear correlation between the presence of soluble periplasmic subunits related to the CISM catalytic subunit architecture and the presence of an outer membrane. In this context, it is perhaps not surprising that evolution has selected against the loss of biosynthetically expensive Mo-*bis*PGD containing enzymes to the bulk milieu.

Unfortunately, existing data and their interpretation do not provide a logical pathway for the evolution of the CISM catalytic subunit architecture. However, it is clear that the FS0 and Mo/W-*bis*PGD containing subunit architecture arose early on in evolution. Evolutionary pathways leading from the catalytic subunit architecture are easier to define. It is clear from our analyses and those of others that there is a clear link between the FdhF formate dehydrogenase, via a gene fusion with a HoxU type subunit, to the Nqo3/NuoG of mitochondrial complex I. Clearly, further efforts will be required to understand the evolution of the seemingly untraceable non-contiguous domain structure of the CISM catalytic subunit. In the case of the FCP architecture, the picture is much clearer, and a pathway leading from a primordial 2[4Fe-4S] cluster-coordinating homodimer to the existing FCP and TCP architectures can be clearly defined.

Acknowledgements

Paul Stothard of the Canadian Bioinformatics Help Desk (CBHD) is thanked for authorship of a PERL script that removes closely related entries from within a sequence dataset. Craig Knox, also of the CBHD, is thanked for writing a PERL script that extracts taxonomies from SWISSPROT and/or TREMBL flatfile entries. The authors would like to thank Milton H. Saier, Jr. (University of California at San Diego) for inspiring this work. We also wish to thank the following for critical reading of the manuscript: Victor Cheng and Kamila Moquin. Special thanks are due to Dean Schieve for installing and maintaining the EMBOSS package used during the course of this work.

This work was funded by the Canadian Institutes of Health Research and the Canada Foundation for Innovation. J.H.W. holds a Canada Research Chair in Membrane Biochemistry.

References

[1] F. Sargent, Constructing the wonders of the bacterial world: biosynthesis of complex enzymes, Microbiology 153 (2007) 633–651.

[2] G. Unden, J. Bongaerts, Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors, Biochim. Biophys. Acta 1320 (1997) 217–234.

[3] R. Gennis, V. Stewart, Respiration, in: F.C. Neidhardt (Ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, vol. 1, ASM Press, Washington, D.C., 1996, pp. 217–261.

[4] D.J. Richardson, Bacterial respiration: a flexible process for a changing environment, Microbiology 146 (2000) 551–571.

[5] M. Jormakka, B. Byrne, S. Iwata, Protonmotive force generation by a redox loop mechanism, FEBS Lett. 545 (2003) 25–30.

[6] M. Jormakka, B. Byrne, S. Iwata, Formate dehydrogenase—a versatile enzyme in changing environments, Curr. Opin. Struck. Biol. 13 (2003) 418–423.

[7] M. Jormakka, D. Richardson, B. Byrne, S. Iwata, Architecture of NarGH reveals a structural classification of Mo-*bis*MGD enzymes, Structure (Camb) 12 (2004) 95–104.

[8] M. Jormakka, S. Tornroth, B. Byrne, S. Iwata, Molecular basis of proton motive force generation: structure of formate dehydrogenase-N, Science 295 (2002) 1863–1868.

[9] M.G. Bertero, R.A. Rothery, M. Palak, C. Hou, D. Lim, F. Blasco, J.H. Weiner, N.C. Strynadka, Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A, Nat. Struct. Biol. 10 (2003) 681–687.

[10] J.H. Weiner, P.T. Bilous, G.M. Shaw, S.P. Lubitz, L. Frost, G.H. Thomas, J.A. Cole, R.J. Turner, A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins, Cell 93 (1998) 93–101.

[11] B.C. Berks, A common export pathway for proteins binding complex redox cofactors? Mol. Microbiol. 22 (1996) 393–414.

[12] L.-F. Wu, B. Ize, A. Chanal, Y. Quentin, G. Fichant, Bacterial twin-arginine signal peptide-dependent protein translocation pathway: evolution and mechanism, J. Mol. Microbiol. Biotechnol. 2 (2000) 179–189.

[13] P. Mitchell, Possible molecular mechanisms of the protonmotive function of cytochrome systems, J. Theor. Biol. 62 (1976) 327–367.

[14] F. Schneider, J. Löwe, R. Huber, H. Schindelin, C. Kisker, J. Knäben, Crystal structure of dimethyl sulfoxide reductase from *Rhodobacter capsulatus* at 1.88 Å resolution, J. Mol. Biol. 263 (1996) 53–69.

[15] H. Schindelin, C. Kisker, J. Hilton, K.V. Rajagopalan, D.C. Rees, Crystal structure of DMSO reductase: redox-linked changes in molybdopterin coordination, Science 272 (1996) 1615–1621.

[16] C. Kisker, H. Schindelin, D. Baas, J. Retey, R.U. Meckenstock, P.M. Kroneck, A structural comparison of molybdenum cofactor-containing enzymes, FEMS Microbiol. Rev. 22 (1998) 503–521.

[17] E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, Acta Crystallogr., D Biol. Crystallogr. 60 (2004) 2256–2268.

[18] F.F. Morpeth, D.H. Boxer, Kinetic analysis of respiratory nitrate reductase from *Escherichia coli* K12, Biochemistry 24 (1985) 40–46.

[19] M. Sabaty, C. Avazeri, D. Pignol, A. Vermeglio, Characterization of the reduction of selenate and tellurite by nitrate reductases, Appl. Environ. Microbiol. 67 (2001) 5122–5126.

[20] C. Avazeri, R.J. Turner, J. Pommier, J.H. Weiner, G. Giordano, A. Vermeglio, Tellurite reductase activity of nitrate reductase is responsible for the basal resistance of *Escherichia coli* to tellurite, Microbiol. 143 (1997) 1181–1189.

[21] T. Satoh, F.N. Kurihara, Purification and properties of dimethylsulfoxide reductase containing a molybdenum cofactor from a photodenitrifier, *Rhodopseudomonas sphaeroides* f.s. *denitrificans*, J. Biochem. (Tokyo) 102 (1987) 191–197.

[22] S.P. Hanlon, D.L. Graham, P.J. Hogan, R.A. Holt, C.D. Reeve, A.L. Shaw, A.G. McEwan, Asymmetric reduction of racemic sulfoxides by dimethyl sulfoxide reductases from *Rhodobacter capsulatus*, *Escherichia* coli and *Proteus* species, Microbiol. 144 (1998) 2247–2253.

[23] H.R. Luckarift, H. Dalton, N.D. Sharma, D.R. Boyd, R.A. Holt, Isolation and characterisation of bacterial strains containing enantioselective DMSO reductase activity: application to the kinetic resolution of racemic sulfoxides, Appl. Microbiol. Biotechnol. 65 (2004) 678–685.

[24] J.L. Simala-Grant, J.H. Weiner, Kinetic analysis and substrate specificity of *Escherichia coli* DMSO reductase. Microbiology 142 (1996) 3231–3239.

[25] J.L. Simala-Grant, J.H. Weiner, Modulation of the substrate specificity of *Escherichia coli* dimethylsulfoxide reductase, Eur. J. Biochem. 251 (1998) 510–515.

[26] B.C. Berks, S.J. Ferguson, J.W.B. Moir, D.J. Richardson, Enzymes and associated electron transport systems that catalyze the respiratory reduction of nitrogen oxides and oxyanions, Biochim. Biophys. Acta 1232 (1995) 97–123.
[27] J. Simon, Enzymology and bioenergetics of respiratory nitrite ammonification, FEMS Microbiol. Rev. 26 (2002) 285–309.
[28] R. Welter, L. Yu, C.A. Yu, The effects of nitric oxide on electron transport complexes, Arch. Biochem. Biophys. 331 (1996) 9–14.
[29] J.L. Johnson, K.V. Rajagopalan, Structural and metabolic relationship between the molybdenum cofactor and urothione, Proc. Natl. Acad. Sci. U. S. A. 79 (1982) 6856–6860.
[30] J.L. Johnson, N.R. Bastian, K.V. Rajagopalan, Molybdopterin guanine dinucleotide: a modified form of molybdopterin identified in the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides* forma specialis *denitrificans*, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 3190–3194.
[31] J.C. Hilton, K.V. Rajagopalan, Identification of the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides* f. sp. *denitrificans* as *bis*(molybdopterin guanine dinucleotide)molybdenum, Arch. Biochem. Biophys. 325 (1996) 139–143.
[32] P. Arnoux, M. Sabaty, J. Alric, B. Frangioni, B. Guigliarelli, J.M. Adriano, D. Pignol, Structural and redox plasticity in the heterodimeric periplasmic nitrate reductase, Nat. Struct. Biol. 10 (2003) 928–934.
[33] J.M. Dias, M.E. Than, A. Humm, R. Huber, G.P. Bourenkov, H.D. Bartunik, S. Bursakov, J. Calvete, J. Caldeira, C. Carneiro, J.J. Moura, I. Moura, M.J. Romao, Crystal structure of the first dissimilatory nitrate reductase at 1.9 Å solved by MAD methods, Struct. Fold Des. 7 (1999) 65–79.
[34] J.C. Boyington, V.N. Gladyshev, S.V. Khangulov, T.C. Stadtman, P.D. Sun, Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocyteine, and an [4Fe-4S] cluster, Science 275 (1997) 1305–1308.
[35] R. Hille, Molybdenum enzymes containing the pyranopterin cofactor: an overview, Met. Ions Biol. Syst. 39 (2002) 187–226.
[36] J.H. Enemark, C.D. Garner, The coordination chemistry and function of the molybdenum centres of the oxomolybdoenzymes, JBIC, J. Biol. Inorg. Chem. 2 (1997) 817–822.
[37] S.J. Nieter Burgmayer, D.L. Pearsall, S.M. Blaney, E.M. Moore, C. Sauk-Schubert, Redox reactions of the pyranopterin system of the molybdenum cofactor, JBIC, J. Biol. Inorg. Chem. 9 (2004) 59–66.
[38] J.J. Moura, C.D. Brondino, J. Trincao, M.J. Romao, Mo and W *bis*-MGD enzymes: nitrate reductases and formate dehydrogenases, JBIC, J. Biol. Inorg. Chem. 9 (2004) 791–799.
[39] D.P. Kloer, C. Hagel, J. Heider, G.E. Schulz, Crystal structure of ethylbenzene dehydrogenase from *Aromatoleum aromaticum*, Structure 14 (2006) 1377–1388.
[40] C.A. Trieber, R.A. Rothery, J.H. Weiner, Multiple pathways of electron transfer in dimethyl sulfoxide reductase of *Escherichia coli*, J. Biol. Chem. 269 (1994) 7103–7109.
[41] C.A. Trieber, R.A. Rothery, J.H. Weiner, Engineering a novel iron–sulfur cluster into the catalytic subunit of *Escherichia coli* dimethyl-sulfoxide reductase, J. Biol. Chem. 271 (1996) 4620–4626.
[42] A. Magalon, M. Asso, B. Guigliarelli, R.A. Rothery, P. Bertrand, G. Giordano, F. Blasco, Molybdenum cofactor properties and [Fe-S] cluster coordination in *Escherichia coli* nitrate reductase A: investigation by site-directed mutagenesis of the conserved His-50 residue in the NarG subunit, Biochemistry 37 (1998) 7363–7370.
[43] R.A. Rothery, M.G. Bertero, R. Cammack, M. Palak, F. Blasco, N.C. Strynadka, J.H. Weiner, The catalytic subunit of *Escherichia coli* nitrate reductase A contains a novel [4Fe-4S] cluster with a high-spin ground state, Biochemistry 43 (2004) 5324–5333.
[44] V. Augier, B. Guigliarelli, M. Asso, P. Bertrand, C. Frixon, G. Giordano, M. Chippaux, F. Blasco, Site-directed mutagenesis of conserved cysteine residues within the β subunit of *Escherichia coli* nitrate reductase. Physiological, biochemical, and EPR characterization of the mutated enzymes. Biochemistry 32 (1993) 2013–2023.
[45] V. Augier, M. Asso, B. Guigliarelli, C. More, P. Bertrand, C. Santini, F. Blasco, M. Chippaux, G. Giordano, Removal of the high-potential [4Fe-4S] center of the β-subunit from *Escherichia coli* nitrate reductase. Physiological, biochemical, and EPR characterization of site-directed mutated enzymes, Biochemistry 32 (1993) 5099–5108.
[46] B. Guigliarelli, M. Asso, C. More, V. Augier, F. Blasco, J. Pommier, G. Giordano, P. Bertrand, EPR and redox characterization of the iron–sulfur centres in nitrate reductases A and Z from *Escherichia coli*. Evidence of a high and low potential class and their relevance in the electron-transfer mechanism. Eur. J. Biochem. 207 (1992) 61–68.
[47] B. Guigliarelli, J. Guillasssier, C. More, P. Sétif, H. Bottin, P. Bertrand, Structural organization of the iron–sulfur centers in *Synechocystis* 6803 photosystem I. EPR study of oriented thylakoid membranes and analysis of the magnetic interactions, J. Biol. Chem. 268 (1993) 900–908.
[48] B. Guigliarelli, A. Magalon, M. Asso, P. Bertrand, C. Frixon, G. Giordano, F. Blasco, Complete coordination of the four Fe-S centres of the β subunit from *Escherichia coli* nitrate reductase. Physiological, biochemical and EPR characterization of site-directed mutants lacking the highest or lowest potential [4Fe-4S] clusters, Biochemistry 35 (1996) 4828–4836.
[49] A. Magalon, D. Lemesle-Meunier, R.A. Rothery, C. Frixon, J.H. Weiner, F. Blasco, Heme axial ligation by the highly conserved His residues in helix II of cytochrome b_{nr} of *Escherichia coli* nitrate reductase A (NarGHI), J. Biol. Chem. 272 (1997) 25652–25658.
[50] A. Magalon, R.A. Rothery, G. Giordano, F. Blasco, J.H. Weiner, Characterization by electron paramagnetic resonance of the role of the *Escherichia coli* nitrate reductase (NarGHI) iron–sulfur clusters in electron transfer to nitrate and identification of a semiquinone radical intermediate, J. Bacteriol. 179 (1997) 5037–5045.
[51] A. Magalon, R.A. Rothery, D. Lemesle-Meunier, C. Frixon, J.H. Weiner, F. Blasco, Inhibitor binding within the NarI subunit (cytochrome b_{nr}) of *Escherichia coli* nitrate reductase A, J. Biol. Chem. 273 (1998) 10851–10856.
[52] R.A. Rothery, A. Magalon, G. Giordano, B. Guigliarelli, F. Blasco, J.H. Weiner, The molybdenum cofactor of *Escherichia coli* nitrate reductase A (NarGHI): effect of a *mobAB* mutation and interactions with [Fe-S] clusters, J. Biol. Chem. 273 (1998) 7462–7469.
[53] R.A. Rothery, F. Blasco, A. Magalon, M. Asso, J.H. Weiner, The hemes of *Escherichia coli* nitrate reductase A (NarGHI): potentiometric effects of inhibitor binding to NarI, Biochemistry 38 (1999) 12747–12757.
[54] R.A. Rothery, F. Blasco, J.H. Weiner, Electron transfer from heme b_L to the [3Fe-4S] cluster of *Escherichia coli* nitrate reductase A (NarGHI), Biochemistry 40 (2001) 5260–5268.
[55] P. Lanciano, A. Vergnes, S. Grimaldi, B. Guigliarelli, A. Magalon, Biogenesis of a respiratory complex is orchestrated by a single accessory protein, J. Biol. Chem. 282 (2007) 17468–17474.
[56] P. Lanciano, A. Magalon, P. Bertrand, B. Guigliarelli, S. Grimaldi, High-stability semiquinone intermediate in nitrate reductase A (NarGHI) from *Escherichia coli* is located in a quinol oxidation site close to heme b_D, Biochemistry 46 (2007) 5323–5329.
[57] R. Cammack, J.H. Weiner, Electron paramagnetic resonance spectroscopic characterization of dimethyl sulfoxide reductase of *Escherichia coli*, Biochemistry 29 (1990) 8410–8416.
[58] V.W. Cheng, R.A. Rothery, M.G. Bertero, N.C. Strynadka, J.H. Weiner, Investigation of the environment surrounding iron–sulfur cluster 4 of *Escherichia coli* dimethylsulfoxide reductase, Biochemistry 44 (2005) 8068–8077.
[59] R.A. Rothery, J.H. Weiner, Alteration of the iron–sulfur composition of *Escherichia coli* dimethyl sulfoxide reductase by site-directed mutagenesis, Biochemistry 30 (1991) 8296–8305.
[60] R.A. Rothery, J.H. Weiner, Topological characterization of *Escherichia coli* DMSO reductase by electron paramagnetic resonance spectroscopy of an engineered [3Fe-4S] cluster, Biochemistry 32 (1993) 5855–5861.
[61] R.A. Rothery, J.H. Weiner, Interaction of an engineered [3Fe-4S] cluster with a menaquinol binding site of *Escherichia coli* DMSO reductase, Biochemistry 35 (1996) 3247–3257.
[62] R.A. Rothery, C.A. Trieber, J.H. Weiner, Interactions between the molybdenum cofactor and iron–sulfur clusters of *Escherichia coli* dimethylsulfoxide reductase, J. Biol. Chem. 274 (1999) 13002–13009.
[63] C.A. Trieber, R.A. Rothery, J.H. Weiner, Consequences of removal of a molybdenum ligand (DmsA-Ser-176) of *Escherichia coli* dimethyl sulfoxide reductase, J. Biol. Chem. 271 (1996) 27339–27345.

[64] A. Manadori, G. Cecchini, I. Shröder, R.P. Gunsalus, M.T. Werth, M.K. Johnson, [3Fe-4S] to [4Fe-4S] cluster conversion in *Escherichia coli* fumarate reductase by site-directed mutagenesis, Biochemistry 31 (1992) 2703–2712.

[65] J. Breton, B.C. Berks, A. Reilly, A.J. Thomson, S.J. Ferguson, D.J. Richardson, Characterization of the paramagnetic iron-sontaining redox centres of *Thiosphaera pantotropha* periplasmic nitrate reductase, FEBS Lett. 345 (1994) 76–80.

[66] V.N. Gladyshev, J.C. Boyington, S.V. Khangulov, D.A. Grahame, T.C. Stadtman, P.C. Sun, Characterization of crystalline formate dehydrogenase H from *Escherichia coli*. Stabilization, EPR spectroscopy, and preliminary crystollographic analysis, J. Biol. Chem. 271 (1996) 8095–8100.

[67] C.A. McDevitt, P. Hugenholtz, G.R. Hanson, A.G. McEwan, Molecular analysis of dimethyl sulphide dehydrogenase from *Rhodovulum sulfidophilum*: its place in the dimethyl sulphoxide reductase family of microbial molybdopterin-containing enzymes, Mol. Microbiol. 44 (2002) 1575–1587.

[68] C.A. McDevitt, G.R. Hanson, C.J. Noble, M.R. Cheesman, A.G. McEwan, Characterization of the redox centers in dimethyl sulfide dehydrogenase from *Rhodovulum sulfidophilum*, Biochemistry 41 (2002) 15234–15244.

[69] C.C. Moser, T.A. Farid, S.E. Chobot, P.L. Dutton, Electron tunneling chains of mitochondria, Biochim. Biophys. Acta 1757 (2006) 1096–1109.

[70] C.C. Page, C.C. Moser, X. Chen, P.L. Dutton, Natural engineering principles of electron tunneling in biological oxidation–reduction, Nature 402 (1999) 47–52.

[71] C.C. Page, C.C. Moser, P.L. Dutton, Mechanism for electron transfer within and between proteins, Curr. Opin. Chem. Biol. 7 (2003) 551–556.

[72] I. Daizadeh, D.M. Medvedev, A.A. Stuchebrukhov, Electron transfer in ferredoxin: are tunneling pathways evolutionarily conserved? Mol. Biol. Evol. 19 (2002) 406–415.

[73] T. Hettmann, R.A. Siddiqui, J. von Langen, C. Frey, M.J. Romao, S. Diekmann, Mutagenesis study on the role of a lysine residue highly conserved in formate dehydrogenases and periplasmic nitrate reductases, Biochem. Biophys. Res. Commun. 310 (2003) 40–47.

[74] R.J. Turner, A.L. Papish, F. Sargent, Sequence analysis of bacterial redox enzyme maturation proteins (REMPs), Can. J. Microbiol. 50 (2004) 225–238.

[75] C.S. Chan, J.M. Howell, M.L. Workentine, R.J. Turner, Twin-arginine translocase may have a role in the chaperone function of NarJ from *Escherichia coli*, Biochem. Biophys. Res. Commun. 343 (2006) 244–251.

[76] T. Krafft, M. Bokranz, O. Klimmeck, I. Schröder, F. Fahrenholz, E. Kojro, A. Kröger, Cloning and nucleotide sequence of the *psrA* gene of *Wolinella succinogenes* polysulphide reductase, Eur. J. Biochem. 206 (1992) 5456–5463.

[77] F. Blasco, C. Iobbi, G. Giordano, M. Chippaux, V. Bonnefoy, Nitrate reductase of *Escherichia coli*: completion of the nucleotide sequence of the *nar* operon and reassessment of the role of the alpha and beta subunits in iron binding and electron transfer, Molec. Gen. Genet. 218 (1989) 249–256.

[78] J.H. Weiner, R.A. Rothery, D. Sambasivarao, C.A. Trieber, Molecular analysis of dimethylsulfoxide reductase: a complex iron–sulfur molybdoenzyme of *Escherichia coli*, Biochim. Biophys. Acta 1102 (1992) 1–18.

[79] F. Blasco, B. Guigliarelli, A. Magalon, M. Asso, G. Giordano, R.A. Rothery, The coordination and function of the redox centres of the membrane-bound nitrate reductases, Cell. Mol. Life Sci. 58 (2001) 179–193.

[80] K. Fukuyama, Y. Nagahara, T. Tsukihara, Y. Katsube, T. Hase, H. Matsubara, Tertiary structure of *Bacillus thermoproteolyticus* [4Fe-4S] ferredoxin. Evolutionary implications for bacterial ferredoxins, J. Mol. Biol. 199 (1988) 183–193.

[81] D.G. George, L.T. Hunt, L.S. Yeh, W.C. Barker, New perspectives on bacterial ferredoxin evolution, J. Mol. Evol. 22 (1985) 20–31.

[82] R. Cammack, D.S. Patil, J.H. Weiner, Evidence that centre 2 in *Escherichia coli* fumarate reductase is a [4Fe-4S] cluster, Biochim. Biophys. Acta 870 (1986) 545–551.

[83] C. Condon, R. Cammack, D.S. Patil, P. Owen, The succinate dehydrogenase of *Escherichia coli*. Immunochemical resolution and biophysical characterization of a 4-subunit enzyme complex, J. Biol. Chem. 260 (1985) 9427–9434.

[84] V.W. Cheng, E. Ma, Z. Zhao, R.A. Rothery, J.H. Weiner, The iron–sulfur clusters in *Escherichia coli* succinate dehydrogenase direct electron flow, J. Biol. Chem. (2006) Electronic publication ahead of print.

[85] J.M. Hudson, K. Heffron, V. Kotlyar, Y. Sher, E. Maklashina, G. Cecchini, F.A. Armstrong, Electron transfer and catalytic control by the iron–sulfur clusters in a respiratory enzyme, *E. coli* fumarate reductase, J. Am. Chem. Soc. 127 (2005) 6977–6989.

[86] T. Ohnishi, C.C. Moser, C.C. Page, P.L. Dutton, T. Yano, Simple redox-linked proton-transfer design: new insights from structures of quinol-fumarate reductase, Structure 8 (2000) R23–R32.

[87] B.L. Berg, J. Li, J. Heider, V. Stewart, Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12: I. Nucleotide sequence of the *fdnGHI* operon and evidence that opal (UGA) encodes selenocysteine, J. Biol. Chem. 266 (1991) 22380–22385.

[88] L. Hederstedt, Succinate:quinone oxidoreductase in the bacteria *Paracoccus denitrificans* and *Bacillus subtilis*, Biochim. Biophys. Acta 1553 (2002) 74–83.

[89] L.S. Huang, G. Sun, D. Cobessi, A.C. Wang, J.T. Shen, E.Y. Tung, V.E. Anderson, E.A. Berry, 3-Nitropropionic acid is a suicide inhibitor of mitochondrial respiration that, upon oxidation by complex II, forms a covalent adduct with a catalytic base arginine in the active site of the enzyme, J. Biol. Chem. 281 (2006) 5965–5972.

[90] F. Sun, X. Huo, Y. Zhai, A. Wang, J. Xu, D. Su, M. Bartlam, Z. Rao, Crystal structure of mitochondrial respiratory membrane protein complex, Cell 121 (2005) 1043–1057.

[91] V. Yankovskaya, R. Horsefield, S. Törnroth, C. Luna-Chavez, H. Miyoshi, C. Léger, B. Byrne, G. Cecchini, S. Iwata, Architecture of succinate dehydrogenase and reactive oxygen species generation, Science 299 (2003) 700–704.

[92] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[93] M.G. Bertero, R.A. Rothery, N. Boroumand, M. Palak, F. Blasco, N. Ginet, J.H. Weiner, N.C. Strynadka, Structural and biochemical characterization of a quinol binding site of *Escherichia coli* nitrate reductase A, J. Biol. Chem. 280 (2005) 14836–14843.

[94] R.A. Rothery, F. Blasco, A. Magalon, J.H. Weiner, The diheme cytochrome *b* subunit (NarI) of *Escherichia coli* nitrate reductase A (NarGHI): structure, function, and interaction with quinols, J. Mol. Microbiol. Biotechnol. 3 (2001) 273–283.

[95] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.

[96] G.J. Mander, E.C. Duin, D. Linder, K.O. Stetter, R. Hedderich, Purification and characterization of a membrane-bound enzyme complex from the sulfate-reducing archaeon *Archaeoglobus fulgidus* related to heterodisulfide reductase from methanogenic archaea, Eur. J. Biochem. 269 (2002) 1895–1904.

[97] R. Hedderich, O. Klimmek, A. Kröger, R. Dirmeier, M. Keller, K.O. Stetter, Anaerobic respiration with elemental sulfur and with disulfides, FEMS Microbiol. Rev. 22 (1999) 353–381.

[98] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.

[99] S. Moller, M.D. Croning, R. Apweiler, Evaluation of methods for the prediction of membrane spanning regions, Bioinformatics 17 (2001) 646–653.

[100] E.G. Bogsch, F. Sargent, N.R. Stanley, B.C. Berks, C. Robinson, T. Palmer, An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria, J. Biol. Chem. 273 (1998) 18003–18006.

[101] R. Groβ, J. Simon, C.R.D. Lancaster, A. Kröger, Identification of histidine residues in *Wolinella succinogenes* hydrogenase that are essential for menaquinone reduction by H_2, Mol. Microbiol. 30 (1998) 639–646.

[102] R. Groβ, R. Pisa, M. Sanger, C.R. Lancaster, J. Simon, Characterization of the menaquinone reduction site in the diheme cytochrome *b* membrane anchor of *Wolinella succinogenes* NiFe-hydrogenase, J. Biol. Chem. 279 (2004) 274–281.
[103] A. Volbeda, M.-H. Charon, C. Piras, E.C. Hatchikian, M. Frey, J.C. Fontecilla-Camps, Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*, Nature 273 (1995) 580–587.
[104] N.K. Heinzinger, S.Y. Fujimoto, M.A. Clark, M.S. Moreno, E.L. Barrett, Sequence analysis of the *phs* operon in *Salmonella typhymurium* and the contribution of thiosulfate reduction to anaerobic metabolism, J. Bacteriol. 177 (1995) 2813–2820.
[105] J.H. Weiner, G. Shaw, R.J. Turner, C.A. Trieber, The topology of the anchor subunit of dimethyl sulfoxide reductase of *Escherichia coli*, J. Biol. Chem. 268 (1993) 3238–3244.
[106] D. Sambasivarao, D.G. Scraba, C.A. Trieber, J.H. Weiner, Organization of dimethyl sulfoxide reductase in the plasma membrane of *Escherichia coli*, J. Bacteriol. 172 (1990) 5938–5948.
[107] D.J. Richardson, N.J. Watmough, Inorganic nitrogen metabolism in bacteria, Curr. Opin. Chem. Biol. 3 (1999) 207–219.
[108] T.A. Clarke, V. Dennison, H.E. Seward, B. Burlat, J.A. Cole, A.M. Hemmings, D.J. Richardson, Purification and spectropotentiometric characterization of *Escherichia coli* NrfB, a decaheme homodimer that transfers electrons to the decaheme periplasmic nitrite reductase complex, J. Biol. Chem. 279 (2004) 41333–41339.
[109] N.R. Stanley, F. Sargent, G. Buchanan, J. Shi, V. Stewart, T. Palmer, B.C. Berks, Behaviour of topological marker proteins targeted to the Tat protein transport pathway, Mol. Microbiol. 43 (2002) 1005–1021.
[110] D. Sambasivarao, H.A. Dawson, G. Zhang, G. Shaw, J. Hu, J.H. Weiner, Investigation of *Escherichia coli* dimethyl sulfoxide reductase assembly and processing in strains defective for the *sec*-independent protein translocation system membrane targeting and translocation, J. Biol. Chem. 276 (2001) 20167–20174.
[111] R.A. Rothery, N. Kalra, R.J. Turner, J.H. Weiner, Sequence similarity as a predictor of the transmembrane topology of membrane-intrinsic subunits of bacterial respiratory chain enzymes, J. Mol. Microbiol. Biotechnol. 4 (2002) 133–150.
[112] Z. Zhao, J.H. Weiner, Interaction of HOQNO with dimethyl sulfoxide reductase of *Escherichia coli*, J. Biol. Chem. 273 (1998) 20758–20763.
[113] P. Geijer, J.H. Weiner, Glutamate 87 is important for menaquinol binding in DmsC of the DMSO reductase (DmsABC) from *Escherichia coli*, Biochim. Biophys. Acta 1660 (2004) 66–74.
[114] A.P. Hinsley, B.C. Berks, Specificity of respiratory pathways involved in the reduction of sulfur compounds by *Salmonella enterica*, Microbiology 148 (2002) 3631–3638.
[115] F. Baymann, E. Lebrun, M. Brugna, B. Schoepp-Cothenet, M.T. Giudici-Orticoni, W. Nitschke, The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes, Philos. Trans. R. Soc. Lond., B Biol. Sci. 358 (2003) 267–274.
[116] V. Méjean, C. Iobbi-Nivol, M. Lepelletier, M. Chippaux, M. Pascal, TMAO anaerobic respiration in *Escherichia coli*: involvement of the *tor* operon, Mol. Microbiol. 11 (1994) 1169–1179.
[117] M. Czjek, J. Dos Santos, J. Pommier, G. Giordano, V. Méjean, R. Haser, Crystal structure of oxidized trimethylamine-*N*-oxide reductase from *Shewanella massilia* at 2.5 Å resolution, J. Mol. Biol. 284 (1998) 435–447.
[118] A. Brige, D. Leys, T.E. Meyer, M.A. Cusanovich, J.J. Van Beeumen, The 1.25 Å resolution structure of the diheme NapB subunit of soluble nitrate reductase reveals a novel cytochrome *c* fold with a stacked heme arrangement, Biochemistry 41 (2002) 4827–4836.
[119] P.J. Ellis, T. Conrads, R. Hille, P. Kuhn, Crystal structure of the 100 kDa arsenite oxidase from *Alcaligenes faecalis* in two crystal forms at 1.64 Å and 2.03 Å, Structure 9 (2001) 125–132.
[120] H. Raaijmakers, S. Macieira, J.M. Dias, S. Teixeira, S. Bursakov, R. Huber, J.J. Moura, I. Moura, M.J. Romao, Gene sequence and the 1.8 Å crystal structure of the tungsten-containing formate dehydrogenase from *Desulfovibrio gigas*, Structure 10 (2002) 1261–1272.
[121] W.B. Pollock, M. Loutfi, M. Bruschi, B.J. Rapp-Giles, J.D. Wall, G. Voordouw, Cloning, sequencing, and expression of the gene encoding the high-molecular-weight cytochrome *c* from *Desulfovibrio vulgaris* Hildenborough, J. Bacteriol. 173 (1991) 220–228.
[122] M. Rossi, W.B.R. Pollock, M.W. Reij, R.G. Keon, R. Fu, G. Voordouw, The *hmc* operon of *Desulfovibrio vulgaris* subsp. *vulgaris* Hildenborough encodes a potential transmembrane redox protein complex, J. Bacteriol. 175 (1993) 4699–4711.
[123] I.A. Pereira, M.J. Romao, A.V. Xavier, J. LeGall, M. Tiexeira, Electron transfer between hydrogenases and mono- and multiheme cytochromes in *Desulfovibrio* sp. JBIC, J. Biol. Inorg. Chem. 3 (1998) 494–498.
[124] L.A. Sazanov, P. Hinchliffe, Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, Science 311 (2006) 1430–1436.
[125] M. Finel, Organization and evolution of structural elements within complex I, Biochim. Biophys. Acta 1364 (1998) 112–121.
[126] J.M. Odom, H.D. Peck Jr., Localization of dehydrogenases, reductases, and electron transfer components in the sulfate-reducing bacterium *Desulfovibrio gigas*, J. Bacteriol. 147 (1981) 161–169.
[127] A. Costa, M. Tiexeira, J. LeGall, J.J. Moura, I. Moura, Formate dehydrogenase from *Desulfovibrio desulfuricans* ATCC 27774: isolation and spectroscopic characterization of the active sites (heme, iron–sulfur centers and molybdenum), JBIC, J. Biol. Inorg. Chem. 2 (1997) 198–208.
[128] C. Sebban, L. Blanchard, M. Bruschi, F. Guerlesquin, Purification and characterization of the formate dehydrogenase from *Desulfovibrio vulgaris* Hildenborough, FEMS Microbiol. Lett. 133 (1995) 143–149.
[129] P.M. Matias, J. Morais, R. Coelho, M.A. Carrondo, K. Wilson, Z. Dauter, L. Sieker, Cytochrome c_3 from *Desulfovibrio gigas*: crystal structure at 1.8 Å resolution and evidence for a specific calcium-binding site, Protein Sci. 5 (1996) 1342–1354.
[130] M. Boll, B. Schink, A. Messerschmidt, P.M. Kroneck, Novel bacterial molybdenum and tungsten enzymes: three-dimensional structure, spectroscopy, and reaction mechanism, Biol. Chem. 386 (2005) 999–1006.
[131] W. Reichenbecher, B. Schink, Towards the reaction mechanism of pyrogallol–phloroglucinol transhydroxylase of *Pelobacter acidigallici*, Biochim. Biophys. Acta 1430 (1999) 245–253.
[132] A. Messerschmidt, H. Niessen, D. Abt, O. Einsle, B. Schink, P.M. Kroneck, Crystal structure of pyrogallol–phloroglucinol transhydroxylase, an Mo enzyme capable of intermolecular hydroxyl transfer between phenols, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 11571–11576.
[133] D.J. Leahy, W.A. Hendrickson, I. Aukhil, H.P. Erickson, Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein, Science 258 (1992) 987–991.
[134] V. Stewart, Y. Lu, A.J. Darwin, Periplasmic nitrate reductase (NapABC enzyme) supports anaerobic respiration by *Escherichia coli* K-12, J. Bacteriol. 184 (2002) 1314–1323.
[135] L.C. Potter, P. Millington, L. Griffifths, G.H. Thomas, J.A. Cole, Competition between *Escherichia coli* strains expressing a periplasic or a membrane-bound nitrate reductase: does Nap confer a selective advantage during nitrate-limited growth? Biochem. J. 344 (1999) 77–84.
[136] L. Philippot, O. Højberg, Dissimilatory nitrate reductases in bacteria, Biochim. Biophys. Acta 1446 (1999) 1–23.
[137] C. Moreno-Vivián, P. Cabello, M. Martínez-Luque, R. Blasco, F. Castillo, Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases, J. Bacteriol. 181 (1999) 6573–6584.
[138] D.E. Pierson, A. Campbell, Cloning and nucleotide sequence of *bisC*, the structural gene for biotin sulfoxide reductase in *Escherichia coli*, J. Bacteriol. 172 (1990) 2194–2198.
[139] V.V. Pollock, M.J. Barber, Molecular cloning and expression of biotin sulfoxide reductase from *Rhodobacter sphaeroides* forma sp. *denitrificans*, Arch. Biochem. Biophys. 318 (1995) 322–332.
[140] J.F. Stolz, P. Basu, J.M. Santini, R.S. Oremland, Arsenic and selenium in microbial metabolism, Annu. Rev. Microbiol. 60 (2006) 107–130.
[141] R. Hille, The mononuclear molybdenum enzymes, Chem. Rev. 96 (1996) 2757–2816.
[142] J.W. Peters, W.N. Lanzilotta, B.J. Lemon, L.C. Seefeldt, X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 Angstrom resolution, Science 282 (1998) 1853–1858.
[143] S. Löscher, T. Burgdorf, I. Zebger, P. Hildebrandt, H. Dau, B. Friedrich, M. Haumann, Bias from H_2 cleavage to production and coordination

changes at the Ni-Fe active site in the NAD^+-reducing hydrogenase from *Ralstonia eutropha*, Biochemistry 45 (2006) 11658–11665.
[144] P. Hinchliffe, L.A. Sazanov, Organization of iron–sulfur clusters in respiratory complex I, Science 309 (2005) 771–774.
[145] E. Nakamaru-Ogiso, T. Yano, T. Yagi, T. Ohnishi, Characterization of the iron–sulfur cluster N7 (N1c) in the subunit NuoG of the proton-translocating NADH–quinone oxidoreductase from *Escherichia coli*, J. Biol. Chem. 280 (2005) 301–307.
[146] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[147] F. Jeanmougin, J.D. Thompson, M. Gouy, D.G. Higgins, T.J. Gibson, Multiple sequence alignment with Clustal X, Trends Biochem. Sci. 23 (1998) 403–405.
[148] I.M. Fearnley, J.E. Walker, Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins, Biochim. Biophys. Acta 1140 (1992) 105–134.
[149] O. Einsle, A. Messerschmidt, R. Huber, P.M. Kroneck, F. Neese, Mechanism of the six-electron reduction of nitrite to ammonia by cytochrome *c* nitrite reductase, J. Am. Chem. Soc. 124 (2002) 11737–11745.
[150] D.J. Eaves, J. Grove, W. Staudenmann, P. James, R.K. Poole, S.A. White, I. Griffiths, J.A. Cole, Involvement of products of the nrfEFG genes in the covalent attachment of haem c to a novel cysteine-lysine motif in the cytochrome *c*552 nitrite reductase from *Escherichia coli*, Mol. Microbiol. 28 (1998) 205–216.
[151] O. Einsle, P. Stach, A. Messerschmidt, J. Simon, A. Kroger, R. Huber, P.M. Kroneck, Cytochrome *c* nitrite reductase from *Wolinella succinogenes*. Structure at 1.6 Å resolution, inhibitor binding, and heme-packing motifs, J. Biol. Chem. 275 (2000) 39608–39616.
[152] M. Rousset, Y. Montet, B. Guigliarelli, N. Forget, M. Asso, P. Bertrand, J.C. Fontecilla-Camps, E.C. Hatchikian, [3Fe-4S] to [4Fe-4S] cluster conversion in *Desulfovibrio fructosovorans* [Ni-Fe] hydrogenase by site-directed mutagenesis, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 11625–11630.
[153] R. Groβ, J. Simon, A. Kröger, The role of the twin-arginine motif in the signal peptide encoded by the *hydA* gene of the hydrogenase from *Wolinella succinogenes*, Arch. Microbiol. 172 (1999) 227–232.
[154] L. Meek, D.J. Arp, The hydrogenase cytochrome *b* heme ligands of *Azotobacter vinelandii* are required for full H_2 oxidation capability, J. Bacteriol. 182 (2000) 3429–3436.
[155] P.M. Matias, A.V. Coelho, F.M. Valente, D. Placido, J. LeGall, A.V. Xavier, I.A. Pereira, M.A. Carrondo, Sulfate respiration in *Desulfovibrio vulgaris* Hildenborough. Structure of the 16-heme cytochrome *c* HmcA AT 2.5 Å resolution and a view of its role in transmembrane electron transfer, J. Biol. Chem. 277 (2002) 47907–47916.
[156] G. Cecchini, I. Schröder, R.P. Gunsalus, E. Maklashina, Succinate dehydrogenase and fumarate reductase from *Escherichia coli*, Biochim. Biophys. Acta 1553 (2002) 140–157.
[157] L. Hederstedt, T. Ohnishi, Progress in succinate:quinone oxidoreductase research, in: L. Ernster (Ed.), Molecular Mechanisms in Bioenergetics, vol. 23, Elsevier Science Publishing Co. Inc, New York, 1992, pp. 163–197.
[158] R. Villemur, M. Lanthier, R. Beaudet, F. Lepine, The *Desulfitobacterium* genus, FEMS Microbiol. Rev. 30 (2006) 706–733.
[159] T. Madsen, D. Licht, Isolation and characterization of an anaerobic chlorophenol-transforming bacterium, Appl. Environ. Microbiol. 58 (1992) 2874–2878.
[160] T. Lee, T. Tokunaga, A. Suyama, K. Furukawa, Efficient dechlorination of tetrachloroethylene in soil slurry by combined use of an anaerobic *Desulfitobacterium* sp. strain Y-51 and zero-valent iron, J. Biosci. Bioeng. 92 (2001) 453–458.
[161] A. Suyama, R. Iwakiri, K. Kai, T. Tokunaga, N. Sera, K. Furukawa, Isolation and characterization of *Desulfitobacterium sp.* strain Y51 capable of efficient dehalogenation of tetrachloroethene and polychloroethanes, Biosci. Biotechnol. Biochem. 65 (2001) 1474–1481.
[162] M. Lanthier, P. Juteau, F. Lepine, R. Beaudet, R. Villemur, *Desulfitobacterium hafniense* is present in a high proportion within the biofilms of a high-performance pentachlorophenol-degrading, methanogenic fixed-film reactor, Appl. Environ. Microbiol. 71 (2005) 1058–1065.
[163] K. Ueda, M. Ohno, K. Yamamoto, H. Nara, Y. Mori, M. Shimada, M. Hayashi, H. Oida, Y. Terashima, M. Nagata, T. Beppu, Distribution and diversity of symbiotic thermophiles, *Symbiobacterium thermophilum* and related bacteria, in natural environments, Appl. Environ. Microbiol. 67 (2001) 3779–3784.
[164] M. Ohno, H. Shiratori, M.J. Park, Y. Saitoh, Y. Kumon, N. Yamashita, A. Hirata, H. Nishida, K. Ueda, T. Beppu, *Symbiobacterium thermophilum* gen. nov., sp. nov., a symbiotic thermophile that depends on co-culture with a *Bacillus* strain for growth, Int. J. Syst. Evol. Microbiol. 50 (2000) 1829–1832.
[165] P.M. Vignais, B. Billoud, J. Meyer, Classification and phylogeny of hydrogenases, FEMS Microbiol. Rev. 25 (2001) 455–501.
[166] P.M. Vignais, A. Colbeau, Molecular biology of microbial hydrogenases, Curr. Issues Mol. Biol. 6 (2004) 159–188.
[167] J.I. Oh, B. Bowien, Structural analysis of the fds operon encoding the NAD^+-linked formate dehydrogenase of *Ralstonia eutropha*, J. Biol. Chem. 273 (1998) 26349–26360.
[168] T. Friedrich, D. Scheide, The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases, FEBS Lett. 479 (2000) 1–5.
[169] D.J. Richardson, B.C. Berks, D.A. Russell, S. Spiro, C.J. Taylor, Functional, biochemical and genetic diversity of prokaryotic nitrate reductases, Cell. Mol. Life Sci. 58 (2001) 165–178.
[170] J.T. Lin, B.S. Goldman, V. Stewart, Structures of genes nasA and nasB, encoding assimilatory nitrate and nitrite reductases in *Klebsiella pneumoniae* M5al, J. Bacteriol. 175 (1993) 2370–2378.
[171] M. Hensel, A.P. Hinsley, T. Nikolaus, G. Sawers, B.C. Berks, The genetic basis of tetrathionate respiration in *Salmonella typhimurium*, Mol. Microbiol. 32 (1999) 275–287.
[172] A. Sirko, M. Zatyka, E. Sadowy, D. Hulanicka, Sulfate and thiosulfate transport in *Escherichia coli* K-12: evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins, J. Bacteriol. 177 (1995) 4134–4136.
[173] T.E. Colyer, N.M. Kredich, In vitro characterization of constitutive CysB proteins from *Salmonella typhimurium*, Mol. Microbiol. 21 (1996) 247–256.
[174] V. Lira-Ruan, G. Sarath, R.V. Klucas, R. Arredondo-Peter, *In silico* analysis of a flavohemoglobin from *Sinorhizobium meliloti* strain 1021, Microbiol. Res. 158 (2003) 215–227.
[175] R.L. Kerby, S.S. Hong, S.A. Ensign, L.J. Coppoc, P.W. Ludden, G.P. Roberta, Genetic and physiological characterization of the *Rhodospirillum rubrum* carbon monoxide dehydrogenase system, J. Bacteriol. 174 (1992) 5284–5294.
[176] P. Rice, I. Longden, A. Bleasby, EMBOSS: The European Molecular Biology Open Software Suite, Trends Genet. 16 (2000) 276–277.
[177] M. Bruschi, F. Guerlesquin, Structure, function, and evolution of bacterial ferredoxins, FEMS Microbiol. Rev. 54 (1988) 155–176.
[178] H. Beinert, Recent developments in the field of iron–sulfur proteins, FASEB J. 4 (1990) 2483–2491.
[179] B. Darimont, R. Sterner, Sequence, assembly and evolution of a primordial ferredoxin from *Thermotoga maritima*, EMBO J. 13 (1994) 1772–1781.
[180] S. Macedo-Ribeiro, B.M. Martins, P.J. Pereira, G. Buse, R. Huber, T. Soulimane, New insights into the thermostability of bacterial ferredoxins: high-resolution crystal structure of the seven-iron ferredoxin from *Thermus thermophilus*, JBIC, J. Biol. Inorg. Chem. 6 (2001) 663–674.
[181] A. Sery, D. Housset, L. Serre, J. Bonicel, C. Hatchikian, M. Frey, M. Roth, Crystal structure of the ferredoxin I from *Desulfovibrio africanus* at 2.3 Å resolution, Biochemistry 33 (1994) 15408–15417.
[182] Y. Nakajima, T. Fujiwara, Y. Fukumori, Purification and characterization of a [3Fe-4S][4Fe-4S] type ferredoxin from hyperthermophilic archaeon, *Pyrobaculum islandicum*, J. Biochem. (Tokyo) 123 (1998) 521–527.
[183] E.D. Duee, E. Fanchon, J. Vicat, L.C. Sieker, J. Meyer, J.M. Moulis, Refined crystal structure of the 2[4Fe-4S] ferredoxin from *Clostridium acidurici* at 1.84 Å resolution, J. Mol. Biol. 243 (1994) 683–695.
[184] M.K. Chan, S. Mukund, A. Kletzin, M.W. Adams, D.C. Rees, Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase, Science 267 (1995) 1463–1469.